Rev Quant Finan Acc (2008) 30:25–47
DOI 10.1007/s11156-007-0042-y

ORIGINAL PAPER

Herding, momentum and investor over-reaction

Rani Hoitash · Murugappa (Murgie) Krishnan

Published online: 12 September 2007

Abstract In this paper we study the impact of noise or quality of prices on returns. The noise arises from herding by market participants beyond what is justified by information. We construct a firm-quarter-specific measure of speculative intensity (SPEC) based on autocorrelation in daily trading volume adjusted for the amount of information available, and find that speculative intensity has a significant positive impact on returns. Both cross-sectional and time series variation in SPEC are consistent with conventional wisdom, and with implications of theories of herding as in DeLong et al. (1990, J Political Econ 98(4):703–738). We find that high-SPEC firms drive the returns to momentum trading strategies and that investor over-reaction is significant only in the case of high-SPEC firms.

Keywords Noise in prices · Measuring speculation · Herding not due to information · Momentum trading · Investor over-reaction

JEL Classifications D83 · D84 · G12 · G14

1 Introduction

Jegadeesh and Titman (1993) document significantly positive returns to a trading rule based on a momentum strategy and demonstrate that the profitability of such trading rules persists over time (Jegadeesh and Titman 2001). DeLong et al. (1990) provide a theory that offers one candidate explanation: they suggest that one source of noise in a financial

Data Availability: The data used in this study is available from public sources cited in the text.

R. Hoitash
Department of Accountancy, Bentley College, 175 Forest Street, Waltham, MA 02452-4705, USA
e-mail: rhoitash@bentley.edu

Murugappa (Murgie) Krishnan (✉)
Department of Accounting, Yeshiva University, 500 West 185th Street, New York, NY 10033, USA
e-mail: murgie@gmail.com

market is speculative intensity or herding by traders beyond what is justified by information about fundamentals. Such herding can give rise to risk beyond what is warranted by fundamentals or information and hence cause prices to deviate from fundamentals. This is one potential explanation for the persistent returns to momentum strategies, and for investor over-reaction. In this paper we construct a firm-quarter specific empirical measure of such speculative intensity or herding (which we call SPEC), and show that it is, as predicted by DeLong et al. (1990), positively associated with absolute returns. We then show that high-SPEC firms can explain momentum returns and that investor over-reaction is significant only for high-SPEC firms.

Conventional wisdom has sometimes suggested that speculative intensity cannot persist, because speculators would incur losses and be arbitraged away. However, recent work has suggested that this deviation of prices from fundamentals can be sustained when rational participants in the market such as institutional traders have short horizons and cannot successfully employ contrarian strategies to outlast speculators.[1] Hence speculative intensity can affect prices for long periods of time. We note at the outset that the term speculation itself has been used in many different ways in the literature. For example, Llorente et al. (2002) refer to "informed speculation" as the behavior of sophisticated participants based on information. Throughout this paper, the term speculation refers (as it does in DeLong et al. 1990), to herding not based on information.

The DeLong et al. (1990) and Dow and Gorton (1994) models suggest that speculative intensity can be characterized by traders who mimic other traders for reasons other than correlated information. This suggests that the trading behavior in one period is related to trading behavior in the next period and can be captured by the autocorrelation in daily trading volume for each stock. But autocorrelation in daily volume can also occur due to information (public or private). We purge the autocorrelation of daily trading volume of such informational effects by regressing it on proxies for the amount of information, such as firm size and the number of firms in an industry, following Bhushan (1989a), and "informed speculation," motivated by Llorente et al. (2002).[2] The residual from this regression measures non-information-based mimicking by traders. We call this variable SPEC and it is our measure of speculative intensity or noise in prices.[3]

We validate our measure of non-information-based herding by relating the cross-sectional and time series patterns of SPEC with other metrics of speculative intensity suggested by prior literature. We find that SPEC is higher in high tech industries than in stable industries like insurance and utilities. Several papers suggest that high tech industries attracted more speculative investors especially during the Internet boom period than did stable industries.[4] Further, the time series of the market-wide aggregate of SPEC is consistent with Shiller (1999)'s Bubble Expectations Index. Since stock market bubbles may be caused by speculative trading not based on information, it is interesting to note that SPEC is correlated with the expectation of a bubble by the survey participants. These two relationships between SPEC and other variables that are indicative of speculative trading give us some confidence that SPEC is a reasonable proxy for non-information-based speculative trading.

[1] see, e.g., DeLong et al. (1990), Dow and Gorton (1994), Grinblatt et al. (1995) Sias (2004).

[2] These information proxies are general proxies and do not distinguish between private and public information.

[3] Hereafter we refer to 'noise in prices' and 'speculative intensity' interchangeably.

[4] See, e.g., Sharma et al. (2004) for a more detailed analysis.

Turning to the relation between SPEC and absolute returns, we show that SPEC is positively related to absolute returns. This supports the theoretical conclusions of DeLong et al. (1990) and Dow and Gorton (1994) and suggests that non-information-based speculative trading impacts prices and is not arbitraged away. Anecdotal evidence suggests that day trading has increased in some periods like the late 1990s and hence has contributed to increased speculative activity in those periods. Chordia et al. (2001) provide evidence that trading costs have generally been falling. This reduces the direct cost of uninformed speculative activity. Further, money flowing into mutual funds has also increased over time. Money managers often have short horizons and this also contributes to increased speculative trading by institutional investors over time. We find that SPEC increases with time suggesting that noise in prices increases over time. We also find that the positive impact of SPEC on returns is robust to an alternative definition of SPEC that allows explicitly for autoregressive errors, and persists even when we distinguish between positive and negative cumulative abnormal returns.

The structure of this paper is as follows. In Sect. 2 we summarize previous work that we draw upon, while in Sect. 3 we describe the construction of our SPEC measure. Section 4 describes the rest of the data and provides descriptive statistics. Section 5 provides support for the DeLong et al. (1990) proposition relating absolute abnormal returns to SPEC, and Sect. 6 has some additional diagnostic tests. Section 7 analyzes the relationship of SPEC to both returns to momentum strategies and investor over-reaction. Section 8 offers concluding remarks.

2 Related work and hypothesis development

Speculative intensity which consists of herding behavior not justified by information plays an important role in financial markets. This proposition relies essentially on ideas in DeLong et al. (1990) and Dow and Gorton (1994). DeLong et al. construct an example in which there is no fundamental uncertainty (so informational issues are moot) yet the presence of irrational noise traders can create risk. Conventional wisdom before DeLong et al. (1990) held that irrationality could not persist, as buying high and selling low would cut into trading profits. But DeLong et al. (1990) show that such irrational noise traders could persist for long periods if rational traders are risk averse and have short horizons. To drive the irrational traders out of the market the rational traders need to adopt contrarian strategies, e.g., buy when the irrational people are selling. Yet if the rational traders have sufficiently short horizons they may have to cash out before their contrarian strategies begin to bear fruit (e.g., if they are buying when the irrational traders are selling, persistent herding by the latter may keep prices low for very long).

Noise traders who follow a herding strategy benefit from the risk that their own participation creates. They make risk averse informed traders less willing to participate in the market so prices are lower and returns higher, and these higher returns are more likely to be earned by the participating noise traders. This suggests a positive association between returns and the level of speculative intensity. DeLong et al. (1990) also show that if the proportion of irrational noise trading is sufficiently high then even the rational traders will find it optimal to try and predict how the irrational traders will behave, rather than seeking to be guided in any way by their information about fundamentals.

Dow and Gorton (1994) further show that agents with short horizons decide on whether to participate in the market based on expectations of subsequent traders following them. It is important to note that the focal point around which the beliefs of participants converge can arise not necessarily from fundamentals like dividends but even from rumor or events affecting the state of the market that have nothing to do with fundamentals. When a trader purchases a stock because he correctly anticipates that others will follow him and also buy stock, the stock price goes up. With sufficiently long horizons the trader will realize his profits if he either holds the stock till it pays off in cash, or if some other trader purchases the shares from him. If the trader has a short horizon, and cannot hold the stock till it pays off, then he would buy the stock in the first place only if he believes that following him will be a chain of traders trading in the same direction spanning the date when the stock will pay off. So he realizes a profit from the follower who pushes the price up further and helps him complete the roundtrip transaction. Smaller the perceived probability of a trader following the first trader, smaller is the probability that the first trader would trade at all. Dow and Gorton (1994) also make the important point, relevant for our work in this paper, that short horizons magnify the impact of transaction costs. The effect of the transaction costs is to make it difficult to diversify away the speculative risk.

Allen et al. (2002) show that noise in prices from a herding strategy by traders is consistent also with a model of *fully rational behavior*. The key technical ingredient in their work is the result that the law of iterated expectations does not apply to average beliefs. So in a setting reminiscent of a beauty contest in which agents seek to predict average beliefs since it is average beliefs that determine prices, they tend to weight public information—whether about fundamentals or about anything else—more than they should. This can create the same herding by traders that occurs in DeLong et al. (1990) even without the assumption of irrationality.

From our perspective the crux of these papers can be summarized as saying that abnormal returns can be expressed as a function of the degree of herding by rational or irrational investors for reasons other than information. This leads us to our primary hypothesis in the null form:

H_{n1}: Speculative intensity does not affect the absolute value of returns.

We must also note that while SPEC is a determinant of noise that captures the idea of mimicking in DeLong et al. (1990) and similar work, it is different from the noise parameters in noisy rational expectations models (e.g., Hellwig 1980; Kyle 1985) where the 'liquidity shock' is an exogenous random shock to aggregate demand. In the Kyle (1985) model, as the liquidity noise (i.e., the variance of noise traders' demand) increases, order flows will become less informative and price-setters will rely less on order flows. So the association between absolute abnormal returns and Kyle's liquidity noise parameter should be negative, and not positive as in the case of SPEC. These two effects are of course not mutually exclusive, and we may find that they both matter. While the liquidity noise parameters are not observable, given availability of trade and quote data (as in the TAQ data base) it is today possible to construct measures of price changes and order flows, and estimate the primitive parameters of a Kyle (1985) or similar model. We leave such comparisons of the contributions of different sources of noise for future research, and focus here only on the impact of SPEC.

SPEC is intended to measure directly at the level of each firm-quarter, the intensity of herding that DeLong et al. (1990) show is related positively to momentum returns. The greater the value of SPEC the more likely it is that a trader participates in the market because of the anticipation that someone else will follow and participate in the same direction. This allows the trader to complete a roundtrip transaction profitably and cash out. The smaller the value of SPEC the less likely there will be such buying or selling waves. This leads to our second hypothesis in the null form:

H_{n2}: SPEC is not positively associated with momentum or short-run returns.

Over the longer term it is in these stocks that we would expect the market to correct and reverse. This reversal should be more pronounced if the market is described well by the DeLong et al. (1990) model, and herding not driven by information is a key driver of prices. So we should observe greater evidence of investor over-reaction in high-SPEC stocks. On the other hand if the herding is driven by information (in particular correlated information leaking slowly over time), while we may see support for momentum reaction by rejecting Hypothesis 2 we would not expect to see any reversal in long-term returns. This leads to our third hypothesis in the null form:

H_{n3}: SPEC is not negatively associated with long-run returns.

In the next section we focus on the measurement of SPEC (other variables used to test the above hypotheses are defined in Sect. 4).

3 Measuring speculative intensity

Speculative intensity in general can take many forms, for example, it could consist of strictly random behavior (like picking stocks by throwing darts). This strategy will not generate any time series correlation in trading behavior, or any systematic pressure on prices. We ignore this source of speculation a priori, as we believe that the measurement error that this introduces is small and does not vary systematically across stocks or over time. At worst in our tests this would create a bias in favor of the null of no relationship between returns and SPEC. Herding or following fads on the other hand creates serial correlation in trading volume, gives rise to order imbalances and consequently requires price-setting market makers to adjust prices in response to such trading pressure. This is the reason we focus on autocorrelation in daily trading volume as a candidate measure of trading pressure arising from speculative intensity or herding behavior. Prior work on fads has used the change in volume as a measure of herding (Bikhchandani and Sharma 2000). Other research on trading volume measure dispersion of beliefs of investors (Bamber 1987; Jain 1988), and focus on a different statistic based on trading volume: market-adjusted abnormal volume.

Autocorrelation in daily trading volume could also result from strictly *contrarian* strategies: buying a lot when most traders appear to be selling, and vice versa. This strategy would however create *less* pressure on prices, whereas speculative intensity via mimicking creates *more* pressure on prices. Further, contrarian strategies would preclude a positive association between autocorrelation in daily trading volume and the absolute value of cumulative abnormal returns, so again in a test of Hypothesis 1 it would create a bias only in favor of the null, making our results more conservative. The financial press has suggested that following fads or herding is often a key factor in trading.

The popular literature on day trading has highlighted how easy it has become for the average person to now participate in chat rooms, access market information, and to place a trade. Chordia et al. (2001) provide evidence that trading costs have generally been falling. This reduces the direct cost of uninformed speculative activity. A mantra for many small uninformed traders has been to follow the behavior of previous traders, in the hope that there are in turn enough traders who will follow them, so that they may quickly complete a roundtrip transaction before the market changes direction. Unlike institutional traders, such day traders by and large only have limited resources to invest in private information acquisition or in gaining access to corporate management. So their primary concern is with minimizing order execution risk within a very short horizon, which also explains why volume begets more volume.

However, the bulk of the market has always been institutional trading. There can be a concern that autocorrelation in daily volume can be caused by institutional traders breaking up their trades and that institutional traders are not representative of the traders who herd for reasons other than information. Why would mimicking or following a fad also be plausible as an equilibrium strategy for institutional investors?

There is direct evidence in Covrig and Ng (2004) to suggest that the level of institutional trading is *inversely* (not directly) related to autocorrelation in daily trading volume. If it is true that institutional traders only trade based on information, and they break up their trades to limit the price impact of their trades (and so reduce returns), this will generate in our empirical design a bias only in favor of the null, that SPEC does *not* affect returns. This makes our actual finding of a positive relationship more *conservative*.

Other research using a different approach than ours suggest that institutions may also herd for reasons other than information (see, e.g., Grinblatt et al. 1995; Li and Yung 2004; Sias 2004; Sharma et al. 2004). Bikhchandani and Sharma (2000) suggest three reasons for institutional herding. First, other investors may have information about the return on the investment and their actions may reveal this information. Second, the incentives provided, to money managers in particular who invest on behalf of other investors, by the compensation schemes and terms of employment may be such that imitation is rewarded. Third, investors may have an intrinsic preference for conformity (Bikchandani and Sharma 2000, p. 3). A recent paper, Frazzini and Lamont (2005), also makes a related empirical argument: that institutional investors are constrained by individual retail investors, and that fund flows within one class of institutional investors, mutual funds, suggest that such "[F]und flows are dumb money."

A key idea is that even fund managers at fairly well established fund families feel some pressure to invest in stocks about which they know relatively little. A typical mutual fund manager is faced with a tradeoff between two different risks. The strategy of following the crowd and investing heavily in glamour or fad stocks, even when her information about fundamentals does not always offer much support for these stocks' prospects, carries the risk that the price rise is a speculative bubble that will burst. However, not following the crowd also means incurring a risk: the risk of being correct but not having one's competence recognized in time because the bubble does not burst within the investor's horizon. Given the importance of relative performance evaluation in the mutual fund industry most mutual fund managers and investors have short horizons. This assumption of short horizons plays a key role in DeLong et al. (1990), Dow and Gorton (1994), and many similar papers in the asset pricing bubbles literature, in generating an incentive for agents to participate in the market only if they expect others to follow on the same side of the market. It also makes more plausible that the optimal choice of agents such as mutual fund managers is to follow the crowd. So the mimicking strategy is important not only for day traders but also for institutional traders. This is the key in justifying our measure of speculative intensity, which is derived from the autocorrelation in daily trading volume.

While this particular type of speculative intensity gives rise to autocorrelation in volume, it is clear that such autocorrelation can arise also for informational reasons. Beaver (1968) and Bamber (1986, 1987) shows that there is a significant volume response to earnings announcements, while Jain (1988) shows there is a volume response to a variety of other public announcements as well. Suppose correlated information about a firm leaks in small doses over a quarter, that by itself could also cause similar autocorrelation. Hence it is important to perform a statistical correction for the possible contribution of informational reasons on the autocorrelation in daily trading volume before using it as a measure of speculative intensity.

The information releases need not all be public or available from sources such as the Dow Jones News Retrieval Service. But previous work has suggested proxies for the amount of information available in general about a firm. The measures are clearly coarse but have been used widely before. See e.g., Bhushan (1989a, b), whose proxies have also been used in recent work such as Hong, Lim and Stein (2001). These include measures of firm size such as market capitalization or total assets, the number of firms in an industry, the number of analysts, or sales or asset growth. So we define SPEC, our measure of mimicking in the following two-step procedure.

First, for each firm i and quarter q we calculate the coefficient of lagged daily trading volume in a first-order autoregression of daily trading volume ($\mathrm{ACC}_{\mathrm{iq}}$). The equation that we use is:

$$Volume_{diq} = \beta_{1iq} Volume_{d-1,iq} + \beta_{2iq} day1 + \beta_{3iq} day2 + \beta_{4iq} day3 + \beta_{5iq} day4 + \beta_{6iq} day5 + u_{iq} \tag{1}$$

where day1 through day5 are day-of-the-week dummy variables,[5] and $ACC_{iq} \equiv \beta_{1iq}$. Previous work supports the use of first order autoregression showing that first-order autocorrelation in daily trading volume is a salient feature of this time series(e.g., Ajinkya and Jain 1989[6] and Llorente et al. 2002).[7] To account for market wide variations in daily trading volume a market-wide measure of daily trading volume (the sum of NYSE and NASDAQ daily volume) is used to scale each daily volume. Each firm should exist in the CRSP database during 1985–2004 and have fiscal quarters that correspond to the calendar quarters. Because there are occasional gross errors in the CRSP database such as a misplaced decimal, to protect against it as a preliminary step within each firm-quarter we deleted raw volume observations that were beyond 5 standard deviations from the mean. We check for missing days before computing the lagged volume; volume data in CRSP is backfilled for earlier periods, and in earlier years there is more missing data. In each quarter there are about 60 trading days, and, we filter out firm-quarters with less than 25[8] days to ensure we have enough observations for each autocorrelation parameter estimate. From Table 1, the overall average of the adjusted R^2 is 0.642 and all coefficients are significant.

We end up with an average of 54.89 degrees of freedom for the above regressions. Using Fisher's z-transform we examine the significance of each firm-quarter autocorrelation coefficient. This entails dividing each of the autocorrelation coefficients by the standard deviation of the distribution of autocorrelations. The resulting z-scores (over the CRSP universe) range from –2.37 to 29.81, with a mean and median of 2.42 and 2.24, respectively. The first and 99th percentiles of the z-scores are –0.64 and 7.18, respectively.

[5] The 5 days of the week dummies serve as the intercept term, inspired by previous work (e.g., Brusa and Liu 2004).

[6] The primary goal in Ajinkya and Jain (1989) is not to identify the time series properties of individual firms' daily trading volumes but to refine the market model for volume used in studies of volume reaction to earnings announcements.

[7] We also investigated the autocorrelation properties of daily trading volume in the firm-quarters that had at least 25 observations. In an overwhelming preponderance of cases the first lag was the most important.

[8] Our results are robust to limiting our sample to autocorrelations that are calculated using 50 or more trading days.

Table 1 Regression of day t volume on day t–1 volume. This table presents the results for the volume autocorrelation regression which was run separately for each company quarter. We regressed the volume at day t on the volume on day t–1 and 5 day of the week dummies and defined β_1 as the autocorrelation coefficient (ACC_{iq}). To account for market wide variations in daily trading volume a market wide measure of was used to adjust each split adjusted daily volume. Missing days (either t or t–1) were dropped from the sample. The average degrees of freedom for each regression is 54.89 days in each quarter (quarters with less then 25 degrees of freedom were dropped from the sample). The overall average of the Adj R^2 over all firm-quarter regressions is 0.642. The average coefficient and the corresponding t-value is presented below for all variables in the regression. For β_1 the median of the t-values was 2.24 and the range was from –2.3 to 29.8. $Volume_{diq} = \beta_{1iq} Volume_{d-1,iq} + \beta_{2iq} day1 + \beta_{3iq} day2 + \beta_{4iq} day3 + \beta_{5iq} day4 + \beta_{6iq} day5 + u_{iq}$

	Coeff.	T-value
$Volume_{t-1}$	0.284838	2.425***
Monday	0.0001571	2.984***
Tuesday	0.0001568	2.929***
Wednesday	0.000156	2.870***
Thursday	0.0001534	2.862***
Friday	0.0001542	2.881***

***Significant at 1% level

So the overwhelming preponderance of z-scores is significantly positive, suggesting a priori a strong possibility that mimicking behavior is important.

We then regress these autocorrelation coefficients on proxies for the amount of information available about each firm, $\ln(MVE_{iq})$, where MVE is the market value of equity at the end of each quarter, and $SIC3_{iq}$, the number of firms in each 3-digit SIC code (Bhushan 1989a) and $PROP_{iq}$, the ratio of the number of days there are forecasts of earnings in each quarter to the total number of days in the quarter for each firm. Analysts have been described as sophisticated users of information by many prior studies (Brown et al. 1987). Further, analysts tend to revise their estimates of earnings in response to (or ahead of) significant information events (see, e.g., Shroff et al. 2003) and the references therein). Hence the proportion of days with analyst forecasts to total number of days in the quarter (PROP) is a proxy for the amount of information that sophisticated users of financial statements possess about a firm during a quarter.

Llorente et al. (2002) consider a setting that recognizes a different non-overlapping set of trading motivations from the ones in DeLong et al. (1990) that motivated our hypotheses. In particular they do not consider herding noise traders but they do consider the possibility of both trading on information and to rebalance portfolios. From their model and argument, they develop an empirical model (Eq. 12 in their paper) of the following form:

$$return_{t+1} = c0 + c1 * return_t + c2 * return_t * volume_t + u_{t+1}$$

where the return and volume variables are daily variables, and the volume variable represents a detrended variable where a 200-day moving average of log-volume is subtracted from log-volume. Given the theory in Llorente et al., the coefficient c2 varies with the proportion of informed trading: the greater the proportion of information-based trading the more positive c2 will be. We compute the same coefficient in the same way as in Llorente et al. (2002) for our sample (for each firm-quarter), and add this variable (denoted by *LLORENTE*) to our set of information proxies. Our measure for speculative intensity or herding due to reasons other than information (SPEC) is given by the residual in the regression model below.

$$ACC_{iq} = \alpha_i + \beta * \ln(MVE_{iq}) + \gamma * SIC3_{iq} + \phi * PROP_{iq} + \delta * LLORENTE_{iq} + SPEC_{iq} \tag{2}$$

In our main result we use OLS regression, however, given the emphasis on autocorrelation in volume in this paper we also allow for first-order serial correlation in the errors, and estimate the following panel data model:

$$\begin{aligned} &ACC_{iq} = \alpha_i + \beta * \ln(MVE_{iq}) + \gamma * SIC3_{iq} + \phi * PROP_{iq} + \delta * LLORENTE_{iq} + SPEC_{iq} \\ &u_{iq} = SPEC_{iq} - \varphi * u_{i,q-1} \\ &\alpha_i \sim N(0, \sigma^2), \ SPEC_{iq} \sim N(0, s_i^2) \end{aligned} \tag{3}$$

So the model represents an attempt to estimate a panel regression, with autoregressive errors and an intercept that can vary by firm, while yet keeping the number of parameters to be estimated manageable.[9] Our key measure of speculative intensity or herding for each firm i and quarter q is the exogenous shock $SPEC_{iq}$. Table 2 provides some descriptive statistics of SPEC computed from Eqs. 2 and 3 above. Allowing for an autoregressive error in (3) marginally reduces the significance of the RHS variables. More pertinently the correlation between the two alternative measures of SPEC is 0.95 with a p-value of less than 0.001.

A priori readers may wonder if there are other candidate measures (at the level of each firm-quarter) for capturing the idea of herding beyond what is justified by correlated information. Why not start with serial correlation in returns rather than in volume? In our context this would lead to an essentially tautological relationship being tested, since the LHS variable is a measure of abnormal returns. Margin costs may seem like another candidate, but not only is the data harder to come by, casual evidence suggests it does not vary much across firms, and may show greater variation based on the financial status of the final investor or of the financial intermediary. For example, it is now well-documented that LTCM was able to use margin trading more extensively and at much lower costs than most other market participants. Abnormal volume has been used to measure an entirely different attribute—dispersion in posterior beliefs of agents—and not mimicking. On the surface it may appear that it is just the 'inverse' of our SPEC measure but this is misleading as SPEC measures similarity in *strategies* and not beliefs. Further, abnormal volume is an indicator of dispersion of beliefs only under the assumption that differences in preferences and opportunities are not significant. In contrast, the motivation for SPEC rests heavily on differences in preferences or opportunities, in particular that some agents have short horizons. The idea of autocorrelation is directly related to the notion of mimicking, and follows from the models of Dow and Gorton (1994) and DeLong et al. (1990). The next section provides further justification.

3.1 Properties of SPEC

SPEC is simply the residual term defined in the regression model of Eq. 2 above. The grand mean (for all firm-quarters in the above regression) is close to zero by construction. So a SPEC value equal to the mean does not mean zero speculation via mimicking. Rather it

[9] The actual implementation used the MIXED procedure in SAS.

Table 2 SPEC models. We report the coefficient estimates for the period of 1985–2004. Panel A present the results of the following OLS regression (SPEC is the error term): $ACC_{iq} = \alpha_i + \beta * \ln(MVE_{iq}) + \gamma * SIC3_{iq} + \phi * PROP_{iq} + \delta * LLORENTE_{iq} + SPEC_{iq}$ (2) whereas in Panel B we allow for first-order serial correlation in the errors, and estimate the following model: $ACC_{iq} = \alpha_i + \beta * \ln(TA_{iq}) + \gamma * SIC3_{iq} + u_{iq} u_{iq} = SPEC_{iq} - \varphi * u_{i,q-1} \alpha_i \sim N(0, \sigma^2),\ SPEC_{iq} \sim N(0, s_i^2)$ where ACC_{iq} is the autocorrelation coefficient of the daily trading volume for each firm-quarter generated from: $Volume_t = \beta_1 Volume_{t-1} + \beta_2 day1 + \beta_3 day2 + \beta_4 day3 + \beta_5 day4 + \beta_6 day5$. $\text{Ln}(MVE_{iq})$ is the natural log of the market value of equity, $SIC3_{iq}$ is the number of firms in each industry, $PROP_{iq}$ is the proportion of days with analyst forecasts to total number of days in the quarter, Llorente is the C2 coefficient from the following equation: $return_{t+1} = c0 + c1 * return_t + c2 * return_t * volume_t + u_{t+1}$ where $volume_t$ represents the detrended variable obtained by subtracting a 200-day moving average of log-volume from log- $volume_t$ and the corresponding residual (SPEC), is the measure of speculative intensity or non-information based trading, in each given firm i-quarter q

	Coeff.	*T*-value	Adj R^2
Panel A: Regression defining the residual SPEC			
Intercept	0.1327	55.46***	
LN(MVE)	0.0196	47.28***	
SIC (3 digit)	0.0000	24.87***	
PROP	0.2473	28.69***	
LLORENTE	0.0281	13.31***	
			0.073
	Coeff.	*T*-value	*Z*-Value
Panel B: SPEC estimation allowing for first-order serial correlation in the errors			
Intercept	0.0002	0.39	
LN(MVE)	0.0213	43.04***	
SIC (3 digit)	0.00004	21.25***	
PROP	0.225	23.60***	
LLORENTE	0.0327	15.80***	
AR(1)	0.175		54.55

***Significant at one sided 1% level

corresponds to the *long run market average level of speculative intensity* (across all firm-quarters over the entire 20-year period from 1985 to 2004). An individual firm-quarter's SPEC is a *market-adjusted number* measuring the abnormal degree of speculative intensity in a particular stock beyond what is justified by information about that stock.

There are clearly several alternative econometric approaches to calculating SPEC, even given the model in (2), which could be more relevant to other applications. Instead of assuming fixed coefficients we could allow coefficients to vary by quarter. In this case the mean value of SPEC across all firms would be zero in each quarter. We could also use out of sample residuals for SPEC. Defining SPEC as deviations from a long-run market average preserves time series variation that also helps us examine reasonableness of our SPEC measure. So we use the strategy defined above.

To gain some confidence that SPEC does indeed capture this abnormal degree of herding by traders in a firm's stock beyond what is justified by the amount of information available about a firm, it is relevant to consider the inputs to the measure (the volume autocorrelation and the various proxies for information) as we have done so far. But it is also important to consider the *outputs* or the measure itself. To do this we first examined the average SPEC over each year and across various SIC codes. This is shown in Table 3, in which 30 industries with the highest and lowest SPEC values are presented. The 15 industries on the

Table 3 Mean SPEC values grouped by four-digit SIC codes. This table presents the mean across firm-quarters of SPEC for selected four-digit SIC codes. (1) SIC codes were selected after calculating the mean of SPEC for each 4 digit SIC code and eliminating industries with less than 300 firm-quarter observations. The 15 industries with the highest and lowest SPEC values are presented in this table. (2) $SPEC_{iq}$ is the residual term calculated from the following regression: $ACC_{iq} = \alpha_i + \beta * \ln(MVE_{iq}) + \gamma * SIC3_{iq} + \phi * PROP_{iq} + \delta * LLORENTE_{iq} + SPEC_{iq}$ (2) ;where: ACC_{iq} is the autocorrelation coefficient from daily trading volume data, $\ln(MVE_{iq})$ is the natural log of the market value of equity, $SIC3_{iq}$ is the number of firms in each industry, $PROP_{iq}$ is the amount of information over each quarter given by the number of analyst forecast revisions in that quarter, Llorente is the C2 coefficient from the following equation: $return_{t+1} = c0 + c1 * return_t + c2 * return_t * volume_t + u_{t+1}$ where $volume_t$ represents the detrended variable obtained by subtracting a 200-day moving average of log-volume from $log\text{-}volume_t$, and the corresponding residual, is the measure of speculative intensity or non-information based trading, in each given firm i-quarter q

Low Speculation				High Speculation			
SIC	SIC Description	*N*	SPEC	SIC	SIC Description	*N*	SPEC
6020	COMMERCIAL BANKS	7,978	–0.06238	5961	CATALOG, MAIL-ORDER HOUSES	469	0.049343
2711	NEWSPAPER:PUBG, PUBG & PRINT	611	–0.05846	3571	ELECTRONIC COMPUTERS	665	0.049508
6331	FIRE, MARINE, CASUALTY INS	2,098	–0.05168	3825	ELEC MEAS & TEST INSTRUMENTS	590	0.050299
6311	LIFE INSURANCE	942	–0.0502	1040	GOLD AND SILVER ORES	317	0.050711
4931	ELECTRIC & OTHER SERV COMB	823	–0.04958	7990	MISC AMUSEMENT & REC SERVICE	632	0.051069
4923	NATURAL GAS TRANSMIS & DISTR	449	–0.049	3559	SPECIAL INDUSTRY MACHY, NEC	958	0.051273
4213	TRUCKING, EXCEPT LOCAL	909	–0.04621	3663	RADIO,TV BROADCAST, COMM EQ	988	0.051957
4924	NATURAL GAS DISTRIBUTION	838	–0.04605	3845	ELECTROMEDICAL APPARATUS	1,033	0.052178
2621	PAPER MILLS	498	–0.04412	6324	HOSPITAL & MEDICAL SVC PLANS	598	0.063104
2670	CONVRT PAPR,PAPRBRD,EX BOXES	444	–0.04198	3674	SEMICONDUCTOR,RELATED DEVICE	2,361	0.063777
6036	SAVINGS INSTN, NOT FED CHART	701	–0.04172	3577	COMPUTER PERIPHERAL EQ, NEC	695	0.065244
4911	ELECTRIC SERVICES	1,499	–0.04132	8731	COML PHYSICAL, BIOLOGCL RESH	318	0.067951
2911	PETROLEUM REFINING	800	–0.0386	3661	TELE & TELEGRAPH APPARATUS	874	0.07357
6035	SAVINGS INSTN,FED CHARTERED	1,950	–0.03405	3576	COMPUTER COMMUNICATION EQUIP	910	0.08048
2631	PAPERBOARD MILLS	381	–0.03119	3572	COMPUTER STORAGE DEVICES	578	0.085818

left side of the table include many stable and regulated industries, which are believed to be a less speculative group e.g., utilities. The 15 industries on the right side of the table are those that had the highest SPEC values and are generally perceived as being more speculative, e.g., hi-tech industries. So there is a reasonable degree of fit between our measure and a priori expectations or conventional wisdom regarding variations in speculative intensity across different industries.

Why should we expect high and low SPEC in the industries listed in Table 3? One contributing factor is that day trading is more concentrated in the relatively high-tech industries listed in Table 3 where we see high SPEC values, and less prevalent in the more stable and regulated industries in that table where we see low SPEC. The evidence in Sharma et al. (2004) is also consistent with this, though their focus is on internet stocks. Further, the results in Frazzini and Lamont (2005) suggest that institutional investors are constrained by retail individual investors.

The closest measure of speculative intensity found in previous research is the Bubble Expectations Index in Shiller (1999). He surveys the attitudes of institutional investors and market experts twice a year, and then constructs an index by aggregating responses to questions dealing with whether an institutional investor feels stock prices are unstable, and will increase only in the short run. This Bubble Expectations Index can be interpreted either as the degree to which the survey respondents feel investors should be cautious following the date of the survey or as the degree to which they feel there has been an unjustifiable price rise in the immediately preceding period. This index is computed twice a year.

To the extent that the Bubble Expectations Index is trying to measure behavior not justified by information its goal is to capture roughly what we try to do with SPEC. However Shiller (1999)'s index is an aggregate measure for the entire market while we calculate SPEC for each firm-quarter. To compare the mean SPEC for all firms for each half-year with the Bubble Expectations Index we first standardize and center both series. The transformed series are depicted in Fig. 1, which clearly reveals a strong positive correlation. The correlation coefficient is 0.66 and has a two-tailed p-value of 0.002. This gives further validation to our use of SPEC as a measure of speculative intensity. What Fig. 1 also shows us is that SPEC is generally increasing over time. We also document this formally in a regression of SPEC on time that we do not report.

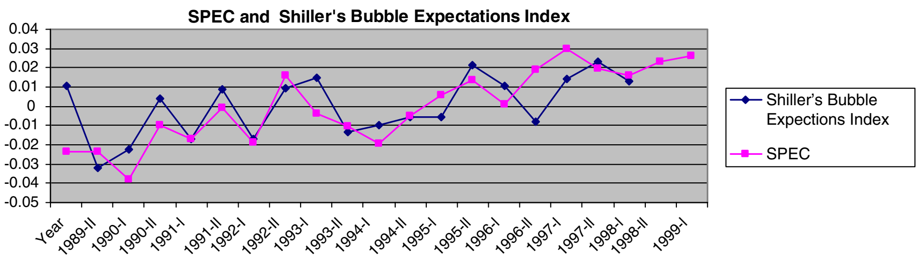

Fig. 1 (1) Shiller's Bubble Expectations Index series is available on semi-annual basis from 1989 to 1998. We computed it by dividing the numbers in Shiller (1999)'s Table 1 by the corresponding standard deviations given in Table 1, and then centering the resulting numbers. (2) SPEC was calculated for corresponding half-years from 1989 to 1999, and standardized and centered in a similar way. (3) The correlation between the two series is 0.69 and has a two-tailed p-value of 0.001

4 Sample, constructed variables and descriptive statistics

4.1 Sample

Our initial sample consists of all the companies in the CRSP daily tapes between the years of 1985–2004.[10] For each company we require that 25 daily trading volume numbers will exist for each quarter, the availability of analyst data from I/B/E/S and the availability of all the required Compustat data items including earnings data. Given these requirements, our final sample includes 102,533 firm-quarter observations.

4.2 Dependent variable

The dependent variable when testing our primary hypothesis is the absolute value of cumulative abnormal returns (ABSCAR). We define the cumulative abnormal returns as the sum of the 3 day abnormal returns around the quarterly earnings announcement. We choose to condition on earnings since public information can also be a source of correlation in daily volume, and earnings are a major example of a firm-specific public announcement. The abnormal return is defined as the firm return less an equally weighted market return from the corresponding size decile calculated over all firms in the CRSP tapes. ABSCAR_{iq} denotes the absolute value of the cumulative abnormal returns (CAR_{iq}).

4.3 Independent Variables

While our primary interest is in SPEC, because we need to control for correlations in daily trading volume arising from common information, we condition on earnings announcements, and use unexpected earnings as a measure of ex post or realized information. Around an earnings announcement it is more plausible that the earnings announcement itself is the primary source of information. Similar to Freeman and Tse (1992) we calculate unexpected earnings (UE) as follows:

$$\mathrm{UE}_{i,q} = (\mathrm{E}_{i,q} - \mathrm{E}_{i,q-4})/\mathrm{P}_{iq}$$

where E_{iq} is the actual quarterly earnings per share before extraordinary items for firm *i* quarter *q*, and P_{iq} is the price per share of firm *i*'s common stock on the last day of quarter *q*. ABSUE_{iq} is the absolute value of unexpected earnings for firm *i* and quarter *q*. SPEC as defined in the previous section is the residual term from Eq. 1.

4.4 Descriptive statistics

Tables 4 and 5 contain descriptive statistics and correlations of our primary variables. All correlations in the table have *p*-values less than 0.001. Examination of SPEC reveals that SPEC values range between –0.952 and 0.532 with a mean of 0. They are also significantly positively correlated with the autocorrelation in daily volume (ACC_{iq}) before we control for the amount of information (using firm size and number of firms in a 3 digit SIC). We

[10] Our sample ends with the 3rd quarter of 2004.

Table 4 Descriptive statistics. This table presents the descriptive statistics for our final sample of 102,533 firm-quarter observations. To remain in the sample firms had to have return and volume data from CRSP, the required data items from Compustat, and analyst forecast data from I/B/E/S.CAR_{iq} is the sum of the daily abnormal returns for firm I quarter q for the period beginning two days after the previous quarterly earnings announcement and ending one day after the current announcement. ABSCAR is the absolute value of CAR. EPS is the earnings per share, UE_{iq} is the unexpected earnings and is calculated as $UE_{iq} = (E_{iq} - E_{i,q-4})/P_{i,q-1}$where E_{iq} is the actual quarterly earnings per share for firm i quarter q, and P_{iq}–1 is the price per share of firm i's common stock on the last day of quarter q–1. ABSUE is the absolute value of UE. $SPEC_{iq}$ is the residual term calculated from the following regression: $ACC_{_{iq}} = \alpha_i + \beta * \ln(MVE_{iq}) + \gamma * SIC3_{iq} + \phi * PROP_{iq} + \delta * LLORENTE_{iq} + SPEC_{iq}$ (2) ; Where ACC_{iq} is the autocorrelation coefficient of the daily trading volume for each firm-quarter generated from: $Volume_t = \beta_1 Volume_{t-1} + \beta_2 day1 + \beta_3 day2 + \beta_4 day3 + \beta_5 day4 + \beta_6 day5$**.** $\ln(MVE_{iq})$ is the natural log of the market value of equity, $SIC3_{iq}$ is the number of firms in each industry, $PROP_{iq}$ is the amount of information over each quarter given by the number of analyst forecast revisions in that quarter, lLorente is the C2 coefficient from the following equation: $return_{t+1} = c0 + c1 * return_t + c2 * return_t * volume_t + u_{t+1}$ where $volume_t$ represents the detrended variable obtained by subtracting a 200-day moving average of log-volume from log- $volume_t$ and the corresponding residual, is the measure of speculative intensity or non-information based trading, in each given firm i-quarter q, and VARRET is the variance of daily returns for each company-quarter adjusted for value-weighted market returns

Variable	*N*	Mean	Std Dev	Median	Minimum	Maximum
ACC	102,533	0.285	0.181	0.286	–0.128	0.721
SPEC	102,533	0.000	0.174	–0.003	–0.952	0.532
CAR	102,533	0.004	0.075	0.002	–0.753	0.917
ABSCAR	102,533	0.050	0.056	0.032	0.000	0.917
EPS	102,533	0.318	0.459	0.280	–1.750	2.150
UE	102,533	–0.001	0.028	0.001	–0.206	0.196
ABSUE	102,533	0.014	0.024	0.005	0.000	0.206
VARRET	102,533	0.001	0.002	0.001	0.000	0.177
LN(MVE)	102,533	6.283	1.663	6.198	1.284	13.312
PROP	102,533	0.074	0.080	0.045	0.011	1.109
SIC3	102,533	221.660	286.537	84.000	1.000	1295.000
LLORENTE	102,533	–0.005	0.260	0.004	–8.898	34.655

measure firm size with market value of equity.[11] SPEC values are positively and significantly correlated with both the signed and the absolute values of CAR, and UE.

To the extent that herding causes volatility in prices and returns, we should expect SPEC to be positively correlated with such measures of volatility. Since prices are not stationary and we have data over a very long period we examine the variance of three different return volatility measures: Variance of raw daily returns for each quarter, the variance of daily returns for each quarter adjusted for value-weighted market returns, and the variance of daily returns for each quarter adjusted for equally-weighted market returns.[12] What is most noteworthy in the correlation table is that despite ACC (raw autocorrelation in daily trading volume) and SPEC being significantly positively correlated (correlation coefficient = 0.96), the correction for the amount of information does purge the information aspects of autocorrelation in trading volume. The correlation between VARRET and ACC is much smaller (correlation coefficient = 0.185), than the correlation

[11] We also used total assets, and it made little difference.

[12] We report results using VARRET defined as the variance of the daily returns adjusted for the value weighted market return calculated for each company quarter.

Table 5 Correlation Pearson (upper triangle) Spearman (lower triangle). This table presents the correlation for selected variables in our final sample of 102,533 firm-quarter observations. To remain in the sample firms had to have return and volume data from CRSP, the required data items from Compustat, and analyst forecast data from I/B/E/S. CAR_{iq} is the sum of the daily abnormal returns for firm i quarter q for the period beginning two days after the previous quarterly earnings announcement and ending one day after the current announcement. ABSCAR is the absolute value of CAR. EPS is the earnings per share, UE_{iq} is the unexpected earnings and is calculated as $UE_{iq} = (E_{iq} - E_{i,q-4})/P_{i,q-1}$ where E_{iq} is the actual quarterly earnings per share for firm i quarter q, and P_{iq-1} is the price per share of firm i's common stock on the last day of quarter q–1. ABSUE is the absolute value of UE. $SPEC_{iq}$ is the residual term calculated from the following regression: $ACC_{iq} = \alpha_i + \beta * \ln(MVE_{iq}) + \gamma * SIC3_{iq} + \phi * PROP_{iq} + \delta * LLORENTE_{iq} + SPEC_{iq}$ (2) ;where ACC_{iq} is the autocorrelation coefficient of the daily trading volume for each firm-quarter generated from: $Volume_t = \beta_1 Volume_{t-1} + \beta_2 day1 + \beta_3 day2 + \beta_4 day3 + \beta_5 day4 + \beta_6 day5$. $\ln(MVE_{iq})$ is the natural log of the market value of equity, $SIC3_{iq}$ is the number of firms in each industry, $PROP_{iq}$ is the amount of information over each quarter given by the number of analyst forecast revisions in that quarter, LLORENTE is the C2 coefficient from the following equation: $return_{t+1} = c0 + c1 * return_t + c2 * return_t * volume_t + u_{t+1}$ where $volume_t$ represents the detrended variable obtained by subtracting a 200-day moving average of log-volume from log- $volume_t$, and the corresponding residual, is the measure of speculative intensity or non-information based trading, in each given firm i-quarter q. and VARRET is the variance of daily returns for each company-quarter adjusted for value-weighted market returns

	ACC	SPEC	CAR	ABSCAR	EPS	UE	ABSUE	VARRET	LN(MVE)	PROP	SIC3
ACC	1	0.9629	–0.0032	0.0971	–0.0350	0.0120	0.0278	0.1424	6.282688	0.2143	0.0683
SPEC	0.9631	1	–0.0012	0.1177	–0.1121	0.0053	0.0692	0.1852	1.662703	0.0000	0.0000
CAR	–0.0020	–0.0026	1	0.0943	0.0458	0.0821	–0.0029	–0.0016	101641	–0.0030	0.0118
ABSCAR	0.0946	0.1147	0.0757	1	–0.1927	–0.0076	0.1291	0.2824	0.240788	–0.0139	0.1130
EPS	–0.0414	–0.1230	0.0655	–0.2174	1	0.2631	–0.3145	–0.3419	0	0.1542	–0.1388
UE	0.0134	0.0163	0.1254	0.0071	0.2755	1	–0.1173	–0.0425	–0.01399	–0.0163	0.0138
ABSUE	0.0232	0.0656	0.0106	0.1238	–0.3025	–0.0446	1	0.2430	–0.13071	–0.0253	0.0409
VARRET	0.1854	0.2544	0.0055	0.3714	–0.5663	–0.0377	0.2691	1	0.376972	–0.0616	0.2237
LN(MVE)	0.232792	0.013707	0.002416	–0.13323	0.437197	0.011887	–0.23393	–0.40095	1	0.607308	–0.03439
PROP	0.2285	0.0401	0.0016	–0.0116	0.2078	–0.0426	–0.0550	–0.1322	0.660241	1	–0.0143
SIC3	0.0705	0.0177	–0.0003	0.0556	–0.1363	0.0123	0.0101	0.1751	–0.05137	–0.0461	1

between VARRET and SPEC (correlation coefficient = 0.254). If we think of any auto-correlation in daily volume as a measure of trading pressure, ACC which includes the effect of information besides non-informational trading, has the effect of reducing return volatility while SPEC which isolates the non-informational trading pressure results in increased volatility. Previous work has not tried to disentangle the two effects.

5 Empirical design and results

Our primary hypothesis is inspired by DeLong et al. (1990): speculative intensity is positively associated with returns. To test this we estimate the following mean centered regression equation:

$$\mathrm{ABSCAR_{iq}} = \beta_1 * \mathrm{ABSUE_{iq}} + \beta_2 * \mathrm{SPEC_{iq}} + \beta_3 * \mathrm{VARRET_{iq}} + \varepsilon_{\mathrm{i}} \tag{4}$$

We need to use absolute values for Cumulative Abnormal Returns and Unexpected Earnings because while SPEC captures speculative intensity beyond what is justified by information, it does not distinguish between buying pressure and selling pressure generated by speculative intensity. For the case of buying pressure the price movement is positively related to trading pressure; for selling pressure, negatively related. Hence we also partition by the sign of CAR, and test if for positive (negative) CAR is positively (negatively) related to SPEC. This is an attempt to separate the buying and selling pressure on increasing and decreasing returns, respectively. If buying pressure is related to increasing returns then SPEC will be positively related to CAR for positive CAR. Further, if selling pressure is related to negative returns, then SPEC will be negatively related to CAR for negative CAR. Since SPEC is unrelated to information we do not partition by the sign of the unexpected earnings. In the above regression we control for the effect of information on returns by way of absolute unexpected earnings. We include the variance of returns to control for change in total risk over time, which could affect the total variability and hence R^2.[13]

The rejection of Hypothesis 1 suggests that SPEC is positively related to $\mathrm{ABSCAR_{iq}}$ and the SPEC Response Coefficient β_2 is positive. Table 6 provides evidence consistent with this. Comparing columns A and B we see that while conditioning for ex post information proxied by Unexpected Earnings does not alter the significance of the SPEC Response Coefficient (which is equal to about 0.021 in both cases), though in keeping with the vast literature on earnings announcements, the Earnings Response Coefficient is also very significant. This adds confidence that our SPEC measure is capturing something *orthogonal* to public information. Partitioning on the sign of CAR and reexamining this hypothesis in each subsample further confirms that SPEC matters. We also explicitly tested for symmetry in the SPEC response coefficient when CAR is positive and when CAR is negative, and find that the response is significantly stronger when CAR is positive. A comparison of Panel A and B in Table 6 shows that our results are qualitatively the same regardless of which of the two measures of SPEC we adopt. The adjusted R-square is also only marginally different. Since we are estimating the equation using panel data, we employ the Fama–MacBeth (1973) methodology and estimate the equation for each quarter

[13] We also ran all of our regressions without VARRET as a control variable, and it turns out the same qualitative conclusions continue to hold. But because a priori it is likely the LHS variance is not constant in a sample spanning 20 years we choose to report the results with VARRET.

Table 6 Regression of signed and absolute cumulative abnormal returns on SPEC, variance of returns (VARRET) and the signed and absolute value of unexpected earnings. We report the coefficient estimates for the following regressions, run for the period of 1985–2004. $ABSCAR_{iq} = \beta_0 + \beta_1 SPEC_{iq} + \beta_3 VAR(\mathrm{Re}turns) + \beta_3 ABSUE\varepsilon_{iq}$ $CAR_{iq} = \beta_0 + \beta_1 SPEC_{iq} + \beta_3 VAR(\mathrm{Re}turns) + \beta_3 UE\varepsilon_{iq}$;where CAR_{iq} is the sum of the three daily abnormal returns for firm i quarter q around the quarterly earnings announcement; $SPEC_{iq}$ is the residual term calculated from the following regression: $ACC_{iq} = \alpha_i + \beta * \ln(MVE_{iq}) + \gamma * SIC3_{iq} + \phi * PROP_{iq} + \delta * LLORENTE_{iq} + SPEC_{iq}$ (2) ; where: ACC_{iq} is the autocorrelation coefficient from daily trading volume data, $\ln(MVE_{iq})$ is the natural log of the market value of equity, $SIC3_{iq}$ is the number of firms in each industry, $PROP_{iq}$ is the proportion of days with analyst forecasts to total number of days in the quarter, LLORENTE is the C2 coefficient from the following equation: $return_{t+1} = c0 + c1 * return_t + c2 * return_t * volume_t + u_{t+1}$ where $volume_t$ represents the detrended variable obtained by subtracting a 200-day moving average of log-volume from log- $volume_t$, and the residual (SPEC), is the measure of speculative intensity or non-information based trading, for each given firm i in quarter q. UE_{iq} is calculated as $UE_{iq} = (E_{iq} - E_{i,q-4})/P_{i,q-1}$ Where E_{iq} is the actual quarterly earnings per share for firm i quarter q, and P_{iq-1} is the price per share of firm i's common stock on the last day of quarter q–1. In Panel C we follow a Fama–Macbeth procedure and report the summary of quarterly estimates using the OLS model

	ABSCAR		Positive CAR	Negative CAR
Panel A: Return regression using SPEC from the OLS model (T-statistics in bold)				
Intercept	0.039	0.038	0.038	–0.039
	192.32***	174.94***	126.49***	–139.69***
SPEC	0.022	0.021	0.172	–0.025
	22.20***	21.72***	12.44***	–18.36***
Variance of returns	8.249	8.73	12.67	–7.021
	88.24***	81.17***	75.04***	–53.34***
Absolute (UE)		0.146		
		20.32***		
UE			0.104	0.101
			12.00***	12.06***
Adj R^2	0.084	0.087	0.109	0.0737
Panel B: Return regression using SPEC from estimation that allows for first-order serial correlation in the errors				
Intercept	0.0393	0.034	0.0377	–0.0393
	191.88***	131.38***	126.22***	–139.34***
SPEC	0.0219	0.0131	0.0172	–0.0254
	22.23***	11.37***	12.45***	–18.39***

Table 6 continued

	ABSCAR		Positive CAR	Negative CAR	
Variance of returns	9.219	13.076	12.639	–6.997	
	87.72***	81.95***	74.64***	–53.01***	
Absolute (UE)		0.1139			
		12.84***			
UE			0.104	0.102	
			11.99***	12.05***	
Adj R^2	0.084	0.108	0.109	0.073	
		Mean	Std Dev	Pr > |*t*|	*t* Value
Panel C: Summary of quarterly regression results using the OLS model					
Intercept		0.033	0.009	<0.0001	31.41***
SPEC		0.013	0.013	<0.0001	8.9***
Variance of returns		13.146	6.143	<0.0001	19.02***
Absolute (UE)		0.098	0.106	<0.0001	8.27***

***Significant at 1% level

and report time series average estimates and statistics in Panel C. It is clear that even after allowing for different coefficient estimates in each period our qualitative conclusion—that ABSCAR is strongly positively related to SPEC—continue to hold.

6 Further analysis and sensitivity checks

We assess robustness of our results in several ways. In defining the initial autocorrelation in daily trading volume, we examine models without day of the week dummies, with and without an intercept and using a log transformed volume. We also construct a measure of volume in which firm volume is not adjusted for market volume. All the autocorrelation specifications are significantly and highly correlated with one another.[14] We also carry these specifications through the resulting values of SPEC to the final returns on SPEC regressions. Our qualitative conclusions still hold. Tables 5 and 6 show that defining SPEC using a model that allows for autoregressive errors still preserves the main qualitative conclusion that SPEC (speculative intensity or non-information based trading) is a key determinant of returns, thus lending support to the theory in DeLong et al. (1990). We replicate the regressions in Table 6 using rank-transformations of the original variables. Again our qualitative conclusions do not change. The adjusted R-square is now 0.15 the coefficient on SPEC is 0.04, and the t-statistic is 14.64 and the rest of the results in that table are significant in a comparable magnitude to the results reported in Table 6.

7 Applications

7.1 Returns to momentum strategies

Jegadeesh and Titman (1993) document that momentum investment strategies yield abnormal returns. The typical strategy is based on rank-ordering stocks on the portfolio formation date based on the returns of the previous 6 months, splitting them into fractiles, and going long in the best performers by this criterion and short in the worst performers. This hedge portfolio is then held for 6 months. Over the years there have been variations on the length of the prior period over which stock performance is examined, the number of fractiles into which all stocks are divided in order to determine the best and worst performers, and the length of the period for which the hedge portfolio is held. But as Jegadeesh and Titman (2001) report the remarkable feature of this trading rule has been its persistent good performance over time. They also survey various explanations for this performance but conclude that a definitive answer still eludes us, and the matter is still an empirical puzzle.

The theory in DeLong et al. (1990) offers a candidate explanation of how momentum returns are generated. Because our SPEC measure is a direct firm-quarter specific measure of the intensity of herding beyond what can be justified by information, it allows us to make a relatively simple test (formally noted in Hypothesis 2) of whether this herding intensity can explain momentum returns.

At each month-end we form hedge portfolios. We classify firms not only based on the previous 6 months' returns but also the average SPEC measure from the previous two

[14] For instance the correlation between the ACCs generated by the firm volumes not adjusted for market volume and the market adjusted volume is 0.78 with p-value of less then 0.0001.

calendar quarters. (Note that some of the variables that enter the SPEC computation are available only on a calendar quarter basis.) We form two sets of ranks based on past returns and SPEC. We define winners and losers as in previous papers based on past return performance, but form three separate hedge portfolios, based on low, medium and high SPEC values. We then examine the performance of the hedge portfolios over the next 6 months.

Panel A in Table 7 summarizes our results. We first examine hedge portfolio returns for a sample that replicates Hong et al. (2001) who also use three categories and obtain numerical estimates and significance values essentially the same as theirs. They dropped the smallest 20% of firms from the CRSP universe, and used a sample from 1980 to 1996. The average return for the top one-third is 1.5% and for the worst performing third is 1%. A hedge portfolio yields a significantly positive return. We do not drop the smallest 20% of firms as they do but construction of our SPEC measure imposes data constraints that also cause small firms to be dropped. We then examine whether the resulting sample also has similar characteristics. It is clear that it does. The top and bottom thirds now yield 1.6% and 1% respectively, and the hedge portfolio returns are again significantly positive. The last 3 columns of the table show the performance of the hedge portfolio for each SPEC category. What is remarkable is that only for high-SPEC do we see significantly positive returns from the hedge portfolio where the top and bottom thirds earn 1.7% and 0.9% returns respectively. This shows that our SPEC measure provides one candidate explanation for momentum returns.

7.2 Investor over-reaction

Where investors over-react in the short run it follows that there will be greater likelihood of a correction and reversal in the longer term, if the argument in DeLong et al. (1990) holds. On the other hand if the herding arose due to the slow leakage of correlated information then we would not see a reversal in the long run. So our strategy for testing Hypothesis 3 is to follow the same portfolio formation strategy as above but examine the returns from months 7 to 36.

Panel B of Table 7 summarizes our results. Again we first examine the CRSP universe and the total sample we have once we impose SPEC-related data constraints. In both cases over months 7 through 36 we have evidence of significant reversal in returns. In the full sample the top and bottom thirds (by past returns) earn 1.3% and 1.5%, respectively. When we look at sub-samples corresponding to low, medium and high SPEC values we find again that only high-SPEC firms have significant larger reversals, with the top and bottom thirds (by past returns) earning 1.2% and 1.7%, respectively, and the hedge portfolio earns significantly negative returns. The evidence of investor over-reaction from these long-run reversals in returns is consistent with the DeLong et al. (1990) argument. Chiao, Cheng and Hung (2005) summarize some prior evidence on over-reaction, and show that evidence from Japan is also consistent with over-reaction after controlling for size and book-to-market.

8 Concluding remarks

In this paper we study the impact of noise or quality of prices on returns. The noise arises from mimicking by market participants beyond what is justified by information. We

Table 7 This table shows the results for the portfolios formed based on six months' lagged raw returns. Panel A shows returns for the first six months, and Panel B the returns from months 7 through 36. The stocks are ranked in ascending order on the basis of their six months lagged returns. Portfolio P1 is an equally weighted portfolio of stocks in the worst performing 33%, portfolio P2 includes the middle 34%, and portfolio P3 includes the best performing 33%. We report the average monthly returns of these portfolios for the entire CRSP sample, our total sample that is limited by the need to compute SPEC, and sub-samples that are based on fractiles of lagged average SPEC. The smallest SPEC value firms are in Group 1 whereas the largest SPEC value firms are in Group 3. Column 2 in Panel A below replicates the results in Hong et al. (2001), who dropped the smallest 20% of firms. We do not but the requirement that SPEC be measured has a very similar effect of dropping the smallest firms

Past returns	Hong et al. (1980–1996)	SPEC sample (1985–2004)	Lagged average spec size		
			1 (low)	2	3 (high)
Panel A: Momentum strategy using raw returns and sorting by SPEC 1985–2004					
P1 (low returns)	0.0108	0.0107	0.0110	0.0119	0.0094
t-Value	2.72***	3.8***	2.4***	2.49***	1.75*
P2 (medium returns)	0.0143	0.0131	0.0133	0.0137	0.0125
t-Value	4.69***	7.19***	4.62***	4.41***	3.56***
P3 (high returns)	0.0158	0.0167	0.0155	0.0168	0.0178
t-Value	4.42***	7.71***	4.87***	4.57***	4.12***
P3–P1	0.0050	0.0059	0.0045	0.0049	0.0084
t-Value	2.96***	3.1***	1.4	1.53	2.38***

Past returns	SPEC sample (1985–2004)	Lagged average spec size		
		1 (low)	2	3 (high)
Panel B: Overreaction strategy using raw returns and sorting by SPEC- 1985–2004				
P1 (low returns)	0.0159	0.0147	0.0159	0.0170
t-Value	6.67***	4.01***	3.93***	3.69***
P2 (medium returns)	0.0141	0.0138	0.0145	0.0141
t-Value	7.88***	4.92***	4.78***	4.07***
P3 (high returns)	0.0134	0.0134	0.0143	0.0125
t-Value	6.42***	4.29***	4.05***	3.03***
P3–P1	–0.0025	–0.0013	–0.0016	–0.0045
t-Value	–3.38***	–1	–1.31	–3.54***

***Significant at one sided 1% level

construct a firm-quarter-specific measure of speculative intensity (SPEC) based on auto-correlation in daily trading volume adjusted for the amount of information available, and find that speculative intensity has a significant positive impact on returns. Both cross-sectional and time series variation in SPEC are consistent with conventional wisdom, and with the implication in DeLong et al. (1990) that in a world where rational information-based traders have short horizons, noise traders who persist in mimicking or herding for reasons unrelated to information cannot be competed away. Such traders not only add volatility to returns but by making other traders more cautious cause current prices to deviate less from the prior mean and so cause subsequent returns to be more extreme.

Our methodological contribution in this paper has been to show how to construct a firm-quarter specific measure of speculative intensity. We believe that this is of independent interest to capital market researchers. This lets us test the explanation advanced in DeLong et al. (1990) in a relatively simple way. We find that high-SPEC firms account for the abnormal returns to a momentum strategy. We also find that high SPEC firms experience the most significant investor over-reaction and consequent reversal in returns over a longer term.

Our SPEC measure is limited by the quality of our proxies for information. We use mainly proxies that have been used widely before. But it is nevertheless good to note that the proxies are coarse, and appropriate caution is therefore needed in interpreting our results.

Our work raises an interesting question for future research: can the publication of a measure like SPEC itself affect market behavior and make traders in a high SPEC environment more cautious? In other words can SPEC serve as an "*amber light*?" Even today it must be the case that when there is a buying wave or a selling wave, market participants seek to distinguish the extent to which this is driven by fundamentals and by noise. Can a coarse quantitative measure like SPEC improve that judgment of market participants?

If publication of such amber signals itself serves to make the behavior of market participants more cautious then the use of circuit breakers and trading halts may fall. There is scope for significant behavioral research in this area. Of course while an amber light makes most drivers cautious and slow down, it does make some drivers aggressive and speed up to beat the light! It is intriguing to consider which tendency would be more dominant in an active financial market.

Acknowledgments We have benefited from comments and suggestions by seminar participants at CUNY-Baruch, Georgetown, Maryland, Rutgers, Waterloo, the Conference on Financial Economics and Accounting, and the Washington Area Finance Association Conference. The paper benefited substantially from help received from Srini Sankaraguruswamy. We also thank Bruce Lehmann, Bikki Jaggi, Prem Jain, Russell Lundholm, Sandeep Patel, Sundaresh Ramnath and Rohan Williamson. Special thanks are due to our discussants at the FEA Conference, Sudhakar Balachandran, and the WAFA Conference, Chris Jones, for very detailed comments.

References

Ajinkya B, Jain PC (1989) The behavior of daily stock market trading volume. J Account Econ 11(4):331–59

Allen F, Morris S, Shin HS (2002) Beauty contests, bubbles and iterated expectations in asset markets. Mimeo, Yale (April 2002) (http://www.econ.yale.edu/~sm326/beauty.pdf)

Bamber LS (1987) Unexpected earnings, firm size and trading volume around quarterly earnings announcements. Account Rev 62(3):510–532

Bamber LS (1986) Information content of annual earnings releases: a trading volume approach. J Account Res 24(1):40–56

Beaver W (1986) The information content of annual earnings announcements. J Account Res (Supplement 1968): 67–92

Bhushan R (1989a) Collection of information about publicly traded firms: theory and evidence. J Account Econ 11(2–3):183–206

Bhushan R (1989b) Firm charecteristics and analyst following. J Account Econ 11(2–3): 255–274

Bikhchandani S, Sharma S (2000) Herd Behavior in Financial Markets: A Review. Working Paper, International Monetary Fund

Brown LD, Griffin P, Hagerman R, Zmijewski M (1987) An evaluation of alternative proxies for market expectations of earnings. J Account Econ 9:159–194

Brusa J, Liu P (2004) The day-of-the-week and the week-of-the-month effects: an analysis of investors' trading activities. Rev Quant Finance Account 23:19–30

Chiao C, Cheng DC, Hung W (2005) Overreaction after controlling for size and book-to-market effects and its mimicking portfolio in Japan. Rev Quant Finance Account 24:65–91

Chordia T, Richard R, Avanidhar S (2001) Market liquidity and trading activity. J Finance 56:501–530
Covrig V, Ng L (2004) Volume autocorrelation, information and investor trading. J Bank Finance 28:2155–2174
DeLong JB, Shleifer A, Summers L, Waldman R (1990) Noise trader risk in financial markets. J Political Econ 98(4):703–738
Dow J, Gorton G (1994) Arbitrage chains. J Finance 49(3):819–849
Fama E, MacBeth J (1973) Risk, return and equilibrium: empirical tests. J Political Econ 81:607–636
Frazzini A, Lamont OA (2005) Dumb money: Mutual fund flows and the cross-section of stock returns, Mimeo, Yale; http://www.nber.org/~confer/2005/URCs05/lamont.pdf
Freeman R, Tse SY (1992) A nonlinear model of security price responses to unexpected earnings. J Account Res 30(2):185–210
Grinblatt M, Titman S, Wermers R (1995) Momentum investment strategies, portfolio performance and herding: a study of mutual fund behavior. Am Econ Rev 85(5):1088–1105
Hellwig MF (1980) On the aggregation of information in competitive markets. J Econ Theory 22:477–498
Hong H, Lim T, Stein JC (2001) Bad news travels slowly: size, analyst coverage and the profitability of momentum strategies. J Finance 55:265–295
Jain PC (1988) Response of hourly stock prices and trading volume to economic news. J Bus 61(2):219–31 (April)
Jegadeesh N, Titman S (1993) Returns to buying winners and selling losers: implications for market efficiency. J Finance 48:65–91
Jegadeesh N, Titman S (2001) Profitability of momentum strategies: an evaluation of alternative explanations. J Finance 56(2):699–720
Kyle AS (1985) Continuous auctions and insider trading. Econometrica 53(6):1315–1335
Li DD, Yung K (2004) Institutional herding in the ADR market. Rev Quant Finance Account 23:5–17
Llorente G, Michaely R, Saar G, Wang J (2002) Dynamic volume-return relation of individual stocks. Rev Financial Studies 15(4):1005–1047
Sharma V, Easterwood JC, Kumar R (2004) Did Institutional Investors Herd into New Economy Stocks? http://www.papers.ssrn.com/sol3/papers.cfm?abstract_id=501423
Shiller RJ (1999) Measuring Bubble Expectations and Investor Confidence. Working Paper, NBER
Shroff PK, Venkatraman R, Xin B (2003) Leaders and followers among security analysts: analysis of impact and accuracy. University of Minnesota, Mimeo
Sias RW (2004) Institutional Herding. Rev Financial Studies 17(1):165–206