THE JOURNAL OF FINANCE • VOL. LIX, NO. 5 • OCTOBER 2004

Correlated Trading and Location

LEI FENG and MARK S. SEASHOLES*

ABSTRACT

This paper analyzes the trading behavior of stock market investors. Purchases and sales are highly correlated when we divide investors geographically. Investors who live near a firm's headquarters react in a similar manner to releases of public information. We are able to make this identification by exploiting a unique feature of individual brokerage accounts in the People's Republic of China. The data allow us to pinpoint an investor's location at the time he or she places a trade. Our results are consistent with a simple, rational expectations model of heterogeneously informed investors.

OVER THE PAST DECADE, financial economists have become increasingly fascinated with the trading decisions of investors. Does a well-defined subset of investors tend to buy or sell the same security *en masse*? What drives this behavior? Are investors choosing their investment strategies by observing the investment decisions of others around them? Are investors driven by some sort of group psychology? Does the presence of a herd mentality affect asset prices?

Alternatively, investors might simply be reacting to the dissemination of information. If subgroups of investors react similarly, financial economists would also measure correlated trading behavior. In this alternative view, it is possible that stock prices and trading patterns are jointly determined as equilibrium outcomes.[1]

This paper addresses the above questions by employing a new data set that is uniquely suited for studying correlated trading. We examine account-level data from the People's Republic of China (or PRC). In the PRC, brokerage rules require that an individual place all of his or her trades through the branch office where he or she opens the account. Individuals are not allowed to open multiple accounts. Our sample consists of trades from seven branch offices. Four are located within the same province of the PRC and three are thousands of kilometers away in another municipality. In addition, our data allow us to

*McKinsey & Co. and University of California, Berkeley, respectively. We thank Ken Froot, Rick Green, Terry Hendershott, Narasimhan Jegadeesh, Terry Odean, Robin Greenwood, Andrei Shleifer, Jeremy Stein, and an anonymous referee for helpful comments. Also, we are grateful to seminar participants at the AEA 2003 Meetings, University of Alberta, Arizona State University, H.E.C Paris, Texas Finance Festival, University of California, Berkeley, and the Utah Winter Finance Conference. Thuy-Uyen Dam provided some excellent research support. Any mistakes are ours alone.

[1] Many empirical studies of herding simply attempt to detect correlated trading and do not consider an equilibrium model of trading and stock returns. Some examples in international markets include Choe, Kho, and Stultz (1999), Kim and Wei (2002a, 2002b), Oehler and Chao (undated), and Lobao and Serra (undated).

track the stock trading behavior of individual investors at a daily frequency. Existing studies typically use monthly or quarterly data.

This PRC brokerage rule is key to our study. It allows us to identify well-defined groups of individuals who are in the same room at the time they are trading.[2] The layout of a typical PRC brokerage office allows for open conservation between investors. Investors congregate in front of an enormous digital display that provides a constant update of stock prices (see Figure 1). Hertz (1998) describes the large amount of social interaction in her ethnographical study of the Shanghai stock exchange. Thus, we expect ex ante to find significant "within-branch" trading patterns.

If investors engage in herd behavior due to animal spirits or group psychology, we would expect the trades of two isolated groups to be uncorrelated. The probability that two isolated groups choose to buy or sell a particular stock on the same day is almost zero. Surprisingly, the first result of this paper shows that isolated groups of investors engage in highly correlated trading behavior. The correlation of net trades (buys minus sells) is positive between two groups in the same region of the country. The correlation of net trades is negative between two groups in different regions of the country. The correlation of total trades (buys plus sells) is positive for any two isolated groups—regardless of location. Thus, there appears to be a lot of evidence of market-wide shocks to trading and little evidence of branch-level (group-psychology) effects.

We then consider the predictions of a rational expectations model with heterogeneously informed investors. The model assumes that investors who live near a firm's headquarters have more precise information about future dividends than investors who live far away. The model predicts that the arrival of public information induces trade as investors at different locations adjust their demands for a risky asset by different amounts. As a result, realized returns should be negatively (positively) correlated with the net trades of the "near" ("far") investors. We show that trading patterns in our data are consistent with these predictions. We also show that net trades of "near" investors load positively on the first principal component of aggregate order flow, while the net trades of "far" investors load negatively on the same factor. Finally, we perform a number of robustness checks.

This paper contributes to our understanding of trading behavior in three main areas. First, we document that the trades of individuals are significantly correlated (contemporaneously), much like the trades of mutual fund managers. The correlation becomes apparent when we condition on location. Second, we

[2] Consider an individual investor who opens an account at the 3 Fuxing Road branch of Huatai Securities in Beijing. This investor is not allowed to place trades at other branches of Huatai Securities (even in Beijing). If the investor travels to Guangdong province, he or she must telephone back to the branch at 3 Fuxing Road in Beijing to place a trade (assuming the investor has phone privileges). If telephoning is not an option, the investor must wait until he or she is back in Beijing before physically going to the office and trading. We concentrate on these physically placed trades, since we know the investor is actually standing in the home branch office when the trade is placed. We know other investors who use the same branch. Therefore, the brokerage rules, combined with the high frequency of our data, allow us to identify isolated groups of investors.

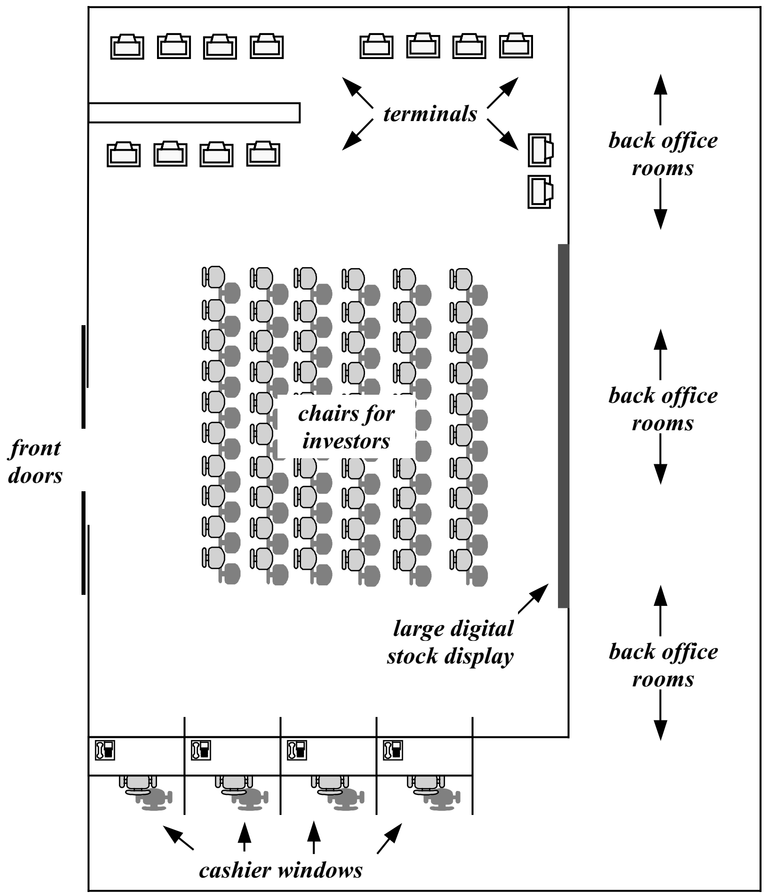

Figure 1. Layout of a typical branch office. This figure presents a schematic drawing of a typical brokerage office. Stock prices are shown on a large, electronic board that covers a good portion of one side of the room. Individual investors can place trades in one of two ways. Some investors place trades through terminals that are located around the edge of the branch office. Investors are able to "log in" by simply swiping a magnetic-strip card at the terminal and then entering a password. Investors enter electronic limit orders. A computer blocks any buy order for which the investor doesn't have sufficient credit or any sell order when the investor does not own the shares. Some margin buying is possible. Other investors place trades at a cashier window after they fill out an order firm. The cashier then enters the buy or sell order into a computer. Again, the computer blocks nonconforming attempts to trade.

show that what has been called "herding" in past studies can be described as trade between asymmetrically informed investors. We outline a simple, rational expectations model and employ an assumption that investors who live near a company's headquarters receive (on average) more precise information about the company than those who live far away. The relationship between the net trades of various groups of investors and stock returns is consistent with predictions of the model. Thus, understanding investors' prior distributions of future payoffs is an important element in understanding trading behavior. Finally, we measure correlated trading behavior at a much higher frequency than most other papers. This allows us to rule out some recently proposed theories of information diffusion.

Brief Review of Related Empirical Work

Before moving on to our actual tests, we briefly review related empirical work.[3] Early studies of herding in financial markets have a hard time detecting either (i) correlated trading or (ii) a relationship between correlated trading and asset returns. Lakonishok, Shleifer, and Vishny (1992) examine the impact of institutional trading in the United States on stock prices. The authors find that if money managers of tax-exempt funds are equally likely to buy or sell a stock in a given quarter, 52.7% of the managers tend to buy (sell) during a quarter while 47.3% do the opposite. This (slight) imbalance of 2.7% could be potentially destabilizing to stock prices, but the authors find little evidence of this. The words "potentially destabilizing" relate to the authors' search for a causal-link from trading behavior to returns. Grinblatt, Titman, and Wermers (1995) find only "weak evidence that funds [tend] to buy and sell the same stocks at the same time."

Recent empirical papers are more successful at detecting correlated trading. Wermers (1999) provides an extensive analysis of the mutual fund industry. Like Lakonishok et al. (1992), he finds more correlated trading in small stocks than in the average stock. He also finds that "stocks that herds buy outperform stocks that they sell by 4% during the following 6 months." The title of his paper, "Mutual Fund Herding and the Impact on Stock Prices," signifies the author's belief that causality runs from trading behavior to asset prices. Kumar and Lee (2002) find a "buy-sell imbalance in individual investors' trades" just like our paper does. However, the authors find the imbalance is correlated with small stock returns at a monthly frequency. They conclude "these findings are broadly consistent with the predictions of a noise trader model in which the systematic activities of individual investors affect the returns of those stocks in which they are concentrated." This interpretation also attributes a causal-link from trades to returns. Finally, in a recent working paper, Barber, Odean, and Zhu (2002) confirm that individual investors (in aggregate) typically have a net trade imbalance.

In other words, is causality running from the purchasing/selling decisions of investors to stock returns? Well-known papers in this area of finance essentially assume this causal-link exists—as is obvious from the papers' titles: "The Impact of Institutional Trading on Stock Prices" by Lakonishok et al. (1992) and "Mutual Fund Herding and the Impact on Stock Prices" by Wermers (1999).

[3] There is a large theoretical literature about herding and information cascades that we will not cover here. For review articles on the entire literature see Devenow and Welch (1996), Hirshleifer and Teoh (2003), and Bikhchandani and Sharma (2000).

There are a few other papers that offer some insights into a possible link between trading behavior and location. Coval and Moskowitz (2001) present one regression that is similar to tests in our paper. The authors regress a herding measure on ownership variables and find "there is a strong inverse relationship between herding activity and geographic proximity." Both our tests and our results differ markedly from theirs, and we discuss these differences later in Section III. Hong, Kubik, and Stein (2002) propose the idea that word-of-mouth communication influences investors' trading decisions. Like this paper does, the authors show that buying/selling is highly correlated within a region. Their data are quarterly and concentrate on mutual fund managers. The authors suggest that an epidemic model explains correlated trading decisions. That is, information diffuses throughout a population. We show that *public* information shocks can explain a large fraction of observed trading behavior, even at frequencies that are too high to allow diffusion. We also discuss these issues in Section III, along with other robustness checks. Finally, Grinblatt and Kelojaru (2001) show Finnish investors are more likely to trade stocks of Finnish firms that are located near them. The authors are also able to examine the marginal effect of language and culture. Unfortunately, our data do not separate language and distance. The authors consider trading volume as measured by the number of buy orders and the number of sell orders. Our paper concentrates on the relation between stock returns and direction of trade (buys minus sells). In this way, the papers provide complementary results.

We now proceed as follows. Section I discusses the data used in this paper; Section II presents our methodology and result; Section III explores alternative hypotheses; and Section IV concludes.

I. Data

We use account-level data to investigate correlated trading in financial markets. Our data come from individual brokerage accounts in the PRC and are uniquely suited for the task at hand. The data represent trades placed between May 4, 1999 and December 4, 2000.

A. Brokerage Accounts in the PRC

Brokerage accounts in the PRC are both similar to, and different from, what we are used to in the United States. A brokerage firm (the firm) has branch offices (branches) throughout the country, region, or city. Many brokerage firms are regionally focused. Individuals open accounts at a branch office and then place all of their trades *through this one branch.* Thus, there is a critical difference in our study between brokerage firms (our data are from one firm) and branch offices (our data come from seven different branches).

A branch office may have a number of ways for investors to place trades: terminals in the branch, cashier windows, telephone service, and computer links. Computer links from private computers are uncommon at this time, effectively leaving three channels with which to place a trade. Consider a brokerage firm with five regional branches in the country's largest cities. An individual who opens an account at the Beijing branch must place all his or her trades with the Beijing branch. Even if the individual visits Shanghai, he or she may not place trades at the local Shanghai branch. Instead, he or she must call Beijing to place a trade (and may only do so if the account has previously been set up to allow phone trades).

B. Individual Investor Data

While investors in the PRC have a number of options for placing trades, we focus on trades that are actually placed at the branch office. We intentionally look at groups of investors who are physically standing near each other during the trading day and, for the time being, do not consider trades that are called-in. We later use telephone trades as a means to recheck our results. We also limit ourselves to secondary trading of shares and do not look at trades relating to IPOs, secondary offerings, or warrants. Our data contain completed trades and not orders that have been submitted and later withdrawn.

Some stocks in the PRC trade infrequently. The highest-volume stock (measured by total value traded in RMB from 1999 to 2000) trades 120.84× more than the lowest-volume stock, 7.96× more than the one-hundredth-ranked stock, and 3.14× more than the seventh-ranked stock. The extreme skewness in trading volume can be seen in Figure 2: a graph of the distribution of the

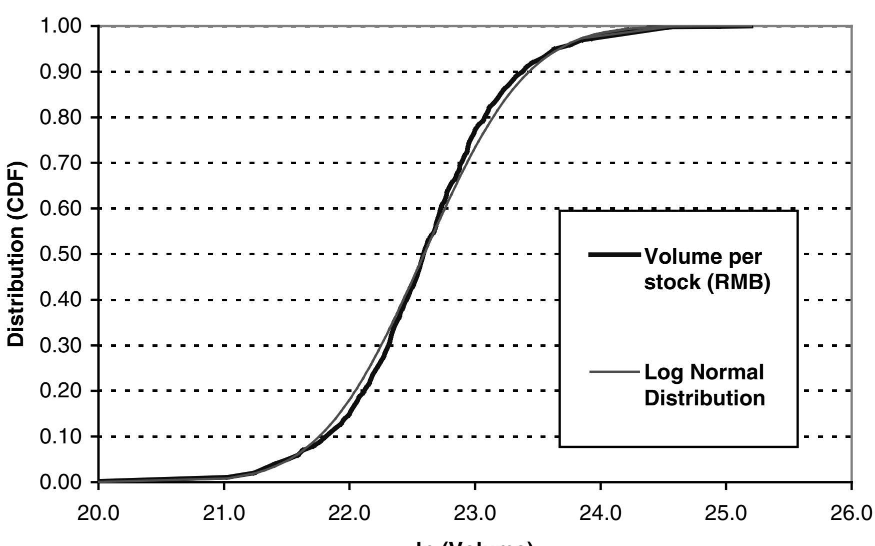

Figure 2. Trading volume in the PRC. The figure graphs the distribution of the natural log of trading volume per listed company. Trading volume is defined as the total value of stock traded (in RMB) for a given company over the 2-year period from 1999 to 2000. The data come from the both the Shenzhen Stock Exchange and the Shanghai Stock Exchange.

natural log of trading volume. Since it is infeasible to measure correlated trading in stocks with low volumes, we choose to look only at trades in active stocks. For simplicity, we limit ourselves to stocks that are listed in Guangdong (on the Shenzhen Stock Exchange) with the company headquarters also in Guangdong. We only look at stocks that are denominated in local currency (RMB). We initially consider the 25 highest-volume stocks as measured by total value traded in 1999 and 2000. We base our selection on total market volume and not in-sample volume. We have rechecked our results using the 100 highest volume stocks and by forming portfolios of stocks. Neither of these twists influences the results in a meaningful way.

Finally, we treat one investor who makes five trades on a given day differently from five different investors making one trade each on the same day. The difference in treatment seems natural when studying correlated trading among investors. We sort our data to include only unique buy and sell orders. That is, if an individual investor makes multiple purchases of a stock on a given day, we count this as one purchase and call it a "unique trade."

C. *Overview of the Data*

Table I presents an overview of the data. We have collected data from seven branches of one brokerage firm. Panel A shows the total trades before we control for investors who break up their trades over the course of 1 day. We can see that the average number of trades per year is 6.03, which is higher than in the United States.[4] Panel B shows only unique trades (defined above) by filtering trades that are broken-up over a day. Panel C shows unique trades that are physically placed in a branch office. The data in Panel C are primarily used throughout this study. The difference between Panel B and Panel C comes from investors who place telephone trades. Table I (Panel C, column (ii)) shows there are 18,575 unique buy-trades placed in branch offices over the 17-month period (80 weeks or 388 trading days) in our sample. The number of sell orders in column (iii) is about 9% less and totals 16,905 trades.

We look at 80 weeks, 25 stocks, and 7 branches, which gives 14,000 groupings (there are 67,900 groupings if we separate trades into 388 trading days instead of 80 weeks). Panel C, column (iv) shows there are 35,480 total unique trades that are physically placed. If trades are spread evenly over time, across stocks, and across branches, we should see an average of 2.53 trades per week in a given stock at a given branch (or 0.52 trades per day-stock-branch). However, trades are not spread evenly over time, across stocks, and across branches. We refer to 2.53 trades per week per stock or 0.52 trades per day per stock as the "data density." A low density is similar to having noisy data and biases estimates of correlated trading to zero. We address this by forming portfolios of stocks, using panel data, or grouping stocks by region. We double-check at all times that our results are consistent regardless of what form we put the data in.

[4] Note: $6.03 = (73,953/7,973) \times (52/80)$, since we have 80 weeks in our sample.

Table I

Overview of Data

This table presents overview statistics of the data used in this study. Data represent stock (equity) trades placed by individual investors in the PRC between May 4, 1999 and December 4, 2000. Trades, or orders, are placed at one of seven brokerage offices that are responsible for maintaining the investors' account data. There are four branches in Guangdong (A, B, C, and D) and three branches in Shanghai (E, F, and G). We concentrate on a sample of buys and sells of high-volume stocks that are listed and headquartered in Guangdong (listed on the Shenzhen Stock Exchange). Panel A shows the total number of trades placed in a particular brokerage branch office during our sample period. We take into account that some individual investors may "break-up" their trades throughout a day. Panel B considers a "unique trade" to be one or multiple trades in the same stock by a single investor on the same day. Panel C considers only trades placed by investors who are physically standing in a particular branch office at the time the trade is placed (by terminal or cashier window).

Branch Office	(i) # of Accts	(ii) # of Buy Orders	(iii) # of Sell Orders	(iv) Buy + Sell Orders
Panel A: Total Trades in Data Set				
A	1,386	8,487	8,141	16,628
B	1,538	7,367	6,504	13,871
C	1,514	8,889	7,878	16,767
D	1,775	7,552	6,803	14,355
E	662	2,901	2,224	5,125
F	751	2,401	1,695	4,096
G	365	1,646	1,465	3,111
Total	7,973	39,243	34,710	73,953
Panel B: Unique Trades in Data Set				
A	1,368	6,003	5,483	11,486
B	1,538	5,316	5,059	10,375
C	1,514	5,887	5,818	11,705
D	1,775	6,348	5,973	12,321
E	662	2,091	1,763	3,854
F	751	1,791	1,346	3,137
G	365	1,183	1,078	2,261
Total	7,973	28,619	26,520	55,139
Panel C: Unique Trades That Were Physically Placed in a Branch Office				
A	957	4,011	3,659	7,670
B	943	3,152	2,943	6,095
C	971	3,851	3,693	7,544
D	963	3,540	3,179	6,719
E	449	1,602	1,377	2,979
F	588	1,490	1,170	2,660
G	263	929	884	1,813
Total	5,134	18,575	16,905	35,480

II. Methodology and Results

We begin our empirical investigation by documenting that investors have highly correlated trades when we condition on location. We then present a simple rational expectations model with a number of testable hypotheses. We show that our data is consistent with the model. Section III offers a number of robustness checks that help reassure the reader that our interpretation of the results is correct.

A. Correlated Trading across Branches

Do investors within an isolated branch behave differently from investors at other, isolated branches? Or, is there a common pattern to the buying and selling behavior of all investors?

To answer these questions, we measure the total trades (buys plus sells) and the net trades (buys minus sells) at each of the seven branches in our sample. We only consider "unique trades" and, for this first test, we form weekly portfolios that consist of all stocks in our sample. Table II presents our initial results. In Panel A, we see the average correlation of total trades between any two

Table II
Regional Trading Correlation

This table presents overview statistics of cross-location (branch) trading activity. Data represent stock (equity) trades placed by individual investors in the PRC between May 4, 1999 and December 4, 2000. This time period represents 80 weeks. Trades are placed at one of seven brokerage branch offices. We concentrate on a sample of buys or sells of high-volume stocks that are listed and headquartered in Guangdong (listed on the Shenzhen Stock Exchange). Panel A: For each branch office, we form a weekly portfolio of unique total trades (buys plus sells) across all stocks in our sample. We then measure the pairwise correlation of these net trades between branches. There are four branches in Guangdong (A, B, C, and D), or six pairs. There are three branches in Shanghai (E, F, and G), or three pairs. There are 12 cross-region pairs. Panel B: We repeat the procedure described above, but use net trades (buys minus sells). We use the asymptotic approximation of the standard error of the correlation coefficient (see text).

	Guangdong	Shanghai
Panel A: Total Trades (Buys + Sells) Average Correlation across Branches		
Guangdong	0.7128	
(z-stat)	(22.22)	
Shanghai	0.4784	0.2512
(z-stat)	(16.86)	(4.15)
Panel B: Net Trades (Buys − Sells) Average Correlation across Branches		
Guangdong	0.3439	
(z-stat)	(8.01)	
Shanghai	−0.2067	0.1457
(z-stat)	(−6.54)	(2.27)

branches is positive. All results are statistically significant at conventional levels.[5]

In Table II (Panel B), we see the average correlation of net trades between branches in Guangdong is positive. In fact, all six correlation-pairs between the four Guangdong branches are positive. The average correlation of net trades between branches in Shanghai is also positive. Again, all three pairs are positive. Notice the average correlation of net trades between a branch in Guangdong and a branch in Shanghai is negative. All 12 pairs are negative. Results are significant at all conventional levels.

Table II presents some strong evidence that group psychology does not determine net buying/selling behavior at the branch level. If group psychology were the main determinant, we would expect the net trades between any two isolated branches to be *uncorrelated*. Clearly, this is not what we see. Instead, there appear to be three important facts. First, Table II (Panel A) shows that trade volume (buys plus sells) is a market-wide phenomenon. There is no evidence that Guangdong investors trade Guangdong stocks amongst themselves. Second, there is clear evidence of correlated trading within a given region of the country. Intraregion correlations are positive between pairs of isolated investor-groups (branch offices). Third, when one region of the country is buying, the other is selling. This third fact is consistent with market-wide shocks to volume. These three facts are not consistent with a group-psychology effect that causes one office to herd on one stock and another office to herd on another stock.

Given the results presented in Table II, we now turn to understanding the economics behind correlated trading.

B. A Simple Rational Expectations Model

In this section, we review the predictions of a noisy, rational expectations model. We then test the implications of such a model with our data. The model is given in Appendix A.[6]

Brief overview of the model: In the model, there are three periods 0, 1, 2 and a continuum of investors who maximize end of period-2 wealth. There is a risky and a riskless asset. The risky asset pays a (random) dividend at the end of period-2. In period-0, investors receive an endowment of the risky asset, a public signal about the dividend of the risky asset, and a private signal about the dividend.

At this point, we make one assumption about the precision of the private signals that investors receive. The precision of the private signal is assumed to

[5] To test for statistical significance, we use the asymptotic approximation of the standard error of the correlation coefficient: $SE(\rho) = \frac{\sqrt{1-\rho^2}}{\sqrt{N-2}}$.

[6] The model in this paper most closely follows Kim and Verrecchia (1991), although the assumption about information asymmetry is most similar to that in Brennan and Cao (1997). We fully recognize that this model is not what is innovative about this paper. The application of information asymmetry to an *intranational* setting (especially to explain correlated trading) is where this paper contributes to the literature.

vary across investors. This assumption is very similar to the one in Brennan and Cao (1997), except we are looking at intranational instead of cross-border (or international) information asymmetry. Investors who live near the headquarters of a traded stock company (the risky asset) are assumed to receive more precise signals (on average) than those investors who live far away from the headquarters. The motivation behind this assumption can be thought of as follows: investors who live near the headquarters may have friends who work for the company, there may be more news stories about the company in the local paper, investors may use the company's products more often, etc. In the United States, it is not hard to imagine that investors read more about their local telephone company than a telephone company across the country.

ASSUMPTION 1 (Location and Information; from Brennan and Cao (1997)): *Investors can be divided into two groups, $i \in \{$"near"," far"$\}$, based on the distance they live from a company's headquarters. Investors who live near the headquarters of a firm receive (on average) more precise information than those who live far away.*

Trade takes place at the end of period-0 so that prices and holding levels are consistent with the investors' beliefs. In period-1, investors receive another public signal about the dividend. Again, they are allowed to trade so the prices and holding levels are consistent with their beliefs. The change in an investor's holdings of the risky asset from period-0 to period-1 represents the net trade of that investor. One-half of the absolute value of the sum of all investors' net trades is the volume or total trade in the risky asset. The change in the equilibrium price from period-0 to period-1 represents the return of the risky asset. These quantities (total trades, net trades, and return on the risky asset) lead to a number of testable hypotheses when combined with the assumption about location and information.

C. Implication Regarding Net Trades (Buys–Sells) and Stock Returns

When positive, public information about the risky asset's dividend is released, all investors raise their (posterior) valuation of the risky asset. However, all investors do not necessarily start with the same priors. By Assumption 1, some investors ("near") have more precise priors than other investors ("far").

When positive information is released, investors with the more diffuse priors ("far") update more than those with more precise priors ("near"). The shift in equilibrium demands for the risky asset causes trade. For a positive signal, "near" investors become net sellers and the "far" investors become net buyers. Again, this happens because the "far" investors update more heavily than the "near" investors (see equation (A7) in Appendix A.3). For a negative signal, the "far" investors also update more heavily. This causes them to become net sellers, while the "near" investors become net buyers.

> H_A (Net trade for "near" investors and returns). If Assumption 1 (location and information) is true, investors who live near the headquarters of a firm have net trades that are negatively correlated with stock returns.

Table III

Regression Analysis of Trading and Stock Returns

This table presents regressions of returns on trading activity. Data represent stock (equity) trades placed by individual investors in the PRC between May 4, 1999 and December 4, 2000. This time period represents 80 trading weeks. Trades are placed at one of seven brokerage offices, and the office is responsible for maintaining the investors' account data. We concentrate on a sample of buys or sells of high-volume stocks that are listed and headquartered in Guangdong (listed on the Shenzhen Stock Exchange). Coefficients are estimated by generalized least squares with an allowance for heteroskedasticity and cross-sectional correlation between stocks. Every week, for every stock, we aggregate net trades in each of the two regions of the country (Guangdong and Shanghai). When then have 2,000 observations for the Guangdong investors and 2,000 observations for the Shanghai investors. The number of observations equals 80 weeks times 25 stocks.

	Investor Location	
	Guangdong (Near)	Shanghai (Far)
Dependent variable	$r_{i,t}/\sigma_i$	$r_{i,t}/\sigma_i$
Independent variable	$Net_{i,t}/\sigma_{Net,i}$	$Net_{i,t}/\sigma_{Net,i}$
$\hat{\beta}$	−0.1220	0.0431
(z-stat)	(−9.58)	(3.44)

COROLLARY TO H_A (net trade for "far" investors and returns): *If Assumption 1 (location and information) is true, investors who live far from the headquarters of a firm have net trades that are positively correlated with stock returns.*

Table III provides a test of Hypothesis H_A and its corollary. Table III offers some of our strongest results. We regress the actual returns on the net trades (buys minus sells) using a generalized least squares procedure that allows for heteroskedasticity and cross-sectional correlation between stocks. Every week, for every stock, we aggregate the net trades in each of the two regions of the country. We see that the net trades of the "near" investors are negatively correlated with returns. We see that the net trades of the "far" investors are positively correlated with returns. Both regression coefficients are significant at conventional levels.

Hypothesis H_A and its corollary are stated in terms of the correlation of net trades and returns. While the regression framework of Table III clearly measures correlation, some readers might appreciate a more direct test. In Table IV, we measure the correlation of net trades and returns. The table shows that the net trades of "near" investors are negatively correlated with returns, while the net trades of "far" investors are positively correlated with returns.

We conclude that our simple rational expectations model, combined with the assumption of location and information, is consistent with our data. That is, trading is the result of portfolio rebalancing. Investors rebalance in response to public information about stocks. The decision by an individual investor to buy or sell (in response to an information shock) is a function of the investor's prior valuation of the company and the new, market-clearing value of the company.

Table IV

Correlation Analysis of Trading and Stock Returns

This table presents the contemporaneous correlation of returns and trading activity. Data represent stock (equity) trades placed by individual investors in the PRC between May 4, 1999 and December 4, 2000. This time period represents eighty trading weeks. Trades are placed at one of seven brokerage offices, and the office is responsible for maintaining the investors' account data. We concentrate on a sample of buys and sells of high-volume stocks that are listed and headquartered in Guangdong (listed on the Shenzhen Stock Exchange). Every week, we form three portfolios: (i) net trades of Guangdong investors across the stocks in our sample; (ii) net trades of Shanghai investors; and (iii) equal-weighted returns across the stocks in our sample. We then measure the correlation of net trades from our two groups of investors with returns. We then have 80 observations for the Guangdong investors and 80 observations for the Shanghai investors. Statistical significance is derived from the asymptotic approximation of the standard error of the correlation coefficient (see text).

	Region	
	Guangdong (Near)	Shanghai (Far)
Correlation coefficient (net trades and returns)	−0.4946	0.2121
(z-stat)	(−5.03)	(1.92)

We now turn to another econometric test that helps solidify our understanding of correlated trading and location.

D. Percent of Trading Explained by Location

We end this section by trying to estimate the fraction of trading behavior that can be explained by location. Why is this important? The regression coefficients in Table III and the correlation coefficients in Table IV give little intuition about the economic importance of our findings. To address this question, we turn to principal component analysis.

In the broadest terms, we can divide trading behavior into two parts: (i) the part that can be explained by location-based effects (common shocks or public news); and (ii) the part that can be explained by other effects (within-branch effects). Examples of within-branch effects have been discussed earlier and include group psychology, insider trading, and white noise.

In order to estimate the fraction of trading that can be explained by location-based effects, we extract a common factor (principal component) from our data. We look at net trades across branches (i.e., across the isolated groups of investors in our sample).

> H_B (Correlated trading across branches). If the decision to buy or sell a given stock is related to market-wide effects, then there is a common component across N-isolated groups of investors. In other words, the first principal component across N-branches explains more than $(1/N)$ of the total variance.

Table V

Principal Component Analysis of Trading Behavior

The table examines the variance of net trading activity. We define net trades as the number of unique buy orders that are physically placed in a branch office, minus the number of unique sell orders. Data come from the PRC between May 4, 1999 and December 4, 2000. This time period represents 80 weeks. We concentrate on a sample of buys or sells of high-volume stocks that are listed and headquartered in Guangdong (listed on the Shenzhen Stock Exchange). This table shows results from a principal component analysis where we have normalized net trades for each stock-branch by its standard deviation. Panel A shows the decision to buy or sell a given stock is correlated across isolated groups of investors. z-stats are calculated using Monte-Carlo methods and the empirical distribution of buys and sells within each branch. Panel B shows loading coefficients that come from a regression of the net trades from a given branch on a constant and the first principal component. Net trades in Guangdong branches load positively on the first principal component, while net trades from Shanghai branches load in the opposite direction on the same principal component. Tests of statistical significance use robust (White) standard errors.

Panel A: Principal Components

	1st Comp	2nd Comp	3rd Comp
Percentage of variance explained	31.8326	15.9632	14.2197
(z-stat)	(7.66)	(5.00)	(3.94)
Cumulative % of variance explained	31.8326	47.7958	62.0155

Panel B: Branch Loadings on the 1st Principal Component

Branch	Loading on 1st component	(z-stat)
A	0.4825	(7.25)
B	0.1838	(3.23)
C	0.2188	(4.70)
D	0.5112	(7.44)
Guangdong average	0.3491	
E	−0.1055	(−2.30)
F	−0.4008	(−6.81)
G	−0.5025	(−3.28)
Shanghai average	−0.3363	

Alternative to H_B (uncorrelated trading across branches): If the decision to buy or sell a given stock is not related to market-wide effects, then there is no common component across N-isolated groups of investors. In other words, the first principal component across N-branches explains ($1/N$) of the total variance, the first two principal components explain ($2/N$) of the variance in total, and so on.

Table V shows strong support for H_B, rejects its alternative, and provides additional support for location-based trading. Table V (Panel A) shows that the first principal component explains 31.83% of the variance of net trades across branches (both Guangdong and Shanghai branches). In the strictest interpretation of our model, we should only consider a single shock (factor) that drives trading and should only measure the first principal component. Thus, we can say that 31.83% of correlated trading can be explained by a single common component (which is public news in the model). The remaining 68.62% can be thought of as within branch effects (including white noise). The first principal component is significant at conventional levels.[7]

We can further bolster our findings by testing whether the first principal component actually represents location-based trading in the way we hypothesize. We regress the net trades from each branch on the first principal component. Table V (Panel B) shows the rather striking results. The Guangdong branches (A, B, C, and D) load heavily on the first component. The Shanghai branches (E, F, and G) load heavily on the second component (but with an opposite sign). Table V (Panel B) is another way to view the results presented earlier in the paper. That is, when investors in Guangdong tend to be buying, investors in Shanghai tend to be selling. There is clear indication of one common factor that affects the entire universe of investors simultaneously (although the effect on net trades depends on location).

To summarize the results up to this point, investors in our sample have correlated trades that become dramatically clear when we condition on location. That is, groups of investors within a region of the country tend to buy and sell together even though the groups (branch offices) are separate and isolated. The trading behavior is consistent with the "near" group of investors having more precise information (than "far" investors) about the prospects of nearby companies. Our interpretation of the results remains constant regardless of which econometric test we use (correlation, regression, or principal component analysis). We now explore a number of alternative hypotheses.

III. Robustness Checks

A. *Implications Regarding Total Trades (Volume) and Stock Returns*

We can use Assumption 1 to formulate specific hypotheses about the sign of the correlation between total trades and stock returns (for details, please see equation (A5) in Appendix A3). The intuition is straightforward. Public news about the payoff of the risky asset causes investors to update their beliefs. When investors update their beliefs and are allowed to trade, volume is correlated with the absolute value of returns. This hypothesis regarding volume and stock returns is by no means unique to our model. It is consistent with our model and serves as a nice alternative check.

H_C (Volume and stock returns). Volume is positively correlated with the absolute value of returns.

The relationship between volume and returns holds for the market as a whole. It also holds for both of our subgroups of investors {"*near*", "*far*"} under

[7] We estimate standard errors using a Monte-Carlo simulation procedure that is explained in Appendix B.

Assumption 1. The reason the relationship holds for our subgroups is that (on average) one group's precision is high and one is low. Thus, neither match the market as a whole.

To test this hypothesis, we regress the absolute value of returns on total trades (buys plus sells). As we did earlier, every week for every stock, we aggregate total trades (buys plus sells) in each of the two regions of the country. We estimate a common coefficient using generalized least squares with an allowance for heteroskedasticity and correlation between stocks. We find the regression coefficients are significantly different from zero at all conventional levels. For "near" investors, the $\hat{\beta}$ coefficient is 0.0642 with a 5.19 z-stat.[8] For the "far" investors, the $\beta - hat$ coefficient is 0.0454 with a 4.47 z-stat. It is clear that the absolute value of returns is positively correlated with trading volume (regardless of whether we look at the volume from "near" or "far" investors). The results are available from the authors upon request and are presented in a format similar to Table III.

B. Turning "Near" and "Far" Around

An obvious and very powerful test of our results can be carried out if we simply flip our definition of "near" and "far." We now concentrate on companies with a headquarters in Shanghai. We still consider stocks listed in Guangdong (on the Shenzhen stock exchange) but now choose a new list of stocks (with a headquarters in Shanghai) and repeat all of the work presented earlier. In particular, we redo the results of Table IV. We calculate the correlation of net trades with returns. For the Guangdong investors (now the "far" investors) the correlation is 0.4904 and significantly positive with a 4.97 z-stat. For the Shanghai investors (now the "near" investors) the correlation is -0.1996 (and significant at the 10% level) with a -1.80 z-stat. The results are available from the authors upon request and are presented in a format similar to Table IV.

C. High-Frequency Results

Another powerful test of our results is to look at frequencies higher than a week. We choose to look at a daily frequency since this is the finest division of stock returns we have available. We measure the correlation of net trades and daily stock returns. We use the full sample of 50 stocks (25 with headquarters in Guangdong and 25 with headquarters in Shanghai). "Near" investors is the union of Guangdong investors trading in Guangdong-headquartered stocks and Shanghai investors trading in Shanghai-headquartered stocks. "Far" investors is the union of Guangdong investors trading in Shanghai-headquarterd stocks and Shanghai investors trading in Guangdong-headquartered stocks.

Trades of the "near" investors are negatively correlated with returns with a -0.0839 correlation coefficient and a -2.34 z-stat. Trades of the "far" investors are positively correlated with returns with a 0.0620 coefficient (and significant at the 10% level) with a 1.73 z-stat. The results are available from the authors upon request and are presented in a format similar to Table IV.

D. Feedback Trading

Papers such as Grinblatt and Keloharju (2000) and Nofsinger and Sias (1999) study feedback-trading strategies of various investor classes. It is possible that we have misidentified the mechanism that appears to cause the contemporaneous correlated trading documented in this paper. It is possible that net trades are actually a response to past price movements. To test if this is the case, we redo Table III, but regress this week's net trades on last week's returns. Appendix C shows our results. We see that neither the "near" investors nor the "far" investors appear to be following a feedback trading strategy. The relationship between net trades and lagged returns is insignificantly different from zero.

In Appendix C (Panel B), we check whether feedback trading might be occurring at a frequency higher than 1 week. We now regress daily net trades (buys minus sells) on lagged returns. Again, we show no significant relationship.

While we would like to look at trading and returns at a higher frequency (intraday), we simply do not have the data to do that. In particular, we do not have tick data from the Shenzhen Stock Exchange (it does not exist). We make the following notes: (i) Our results at a daily frequency represent a 20-times finer (approximately) examination of correlated trading than past studies that use monthly data. (ii) We have shown in Table IV that net trades and contemporaneous returns are correlated at a daily frequency. (iii) If feedback trading exists, there might be a predictable pattern in stock returns (see De Long et al. (1990)). We do not see such a pattern in PRC stock returns. (iv) When examining data at a very high frequency, our definition of correlated trading has to change. With tick-by-tick data, only two trades cross at a time. One trade is a buy and one is a sell. No other trades happen at exactly the same time. Therefore, we might have to look at orders flow (instead of executed trades) to see if groups of investors are posting orders *en masse.* Unfortunately, we do not have these order flow data either.

E. Coval and Moskowitz (2001) Results

As mentioned in our introduction, Coval and Moskowitz (2001) present one regression that is similar to tests in our paper. The authors regress a herding measure on ownership variables and find "there is a strong inverse relationship between herding activity and geographic proximity." Our results are quite different. All trading decisions in our model are determined simultaneously and "herding" does not depend monotonically on location. In fact, we should see net trades ("herding") in one direction by "near" investors and net trades ("herding") in the opposite direction by "far" investors. If we had investors from a "middle" region, we would expect zero net trades.

Table II (Panel B) in our paper shows a higher correlation coefficient among "near" locations than "far" locations. This result is *opposite* to that of Coval and Moskowitz (2001) but only considers Guangdong-headquartered stocks. High-frequency analysis from part C (above) shows roughly similar magnitudes of correlation coefficients when we consider stocks with headquarters in both Guangdong and Shanghai. Note the correlation coefficients are of opposite sign (negative for the "near" locations and positive for the "far" locations). This result is consistent with the intuition given in the paragraph directly above.

F. Epidemic Models and Word-of-Mouth Communication

Economists have long been interested in how information spreads. We believe that investors (primarily) react to public (common) information. A competing view is proposed by Hong et al. (2002) who look at quarterly data and argue that investors "spread information and ideas about stocks directly to one another by word of mouth." The authors suggest an epidemic model to explain their findings. That is, trading decisions over the quarter are correlated, because investors who live near each other pass information between themselves.

While an epidemic model may be consistent with a quarterly observation interval, it may not be consistent with high-frequency data. We reach this conclusion, because we are able to shrink our observation interval from a weekly to a daily frequency. Daily frequency is an important threshold. If word-of-mouth communication plays a part in investors' decisions, then groups of investors would have to meet (talk by phone) each evening (i.e., between trading hours). We do not believe it is likely investors from the same region of a country compare notes after trading hours and spread information amongst themselves. We make this point in part C (above), where we see that net trades within a given region of the country are still significantly correlated with returns (in the direction our model predicts), even at a daily frequency!

We extend the high-frequency tests and look at the *daily* correlation of total trades (buys + sells). We find the market as a whole experiences shocks to total trades at a daily frequency. The correlation of total trades within Guangdong Province is 0.6883 (z-stat is 45.76), the correlation within Shanghai Municipality is 0.2214 (z-stat is 7.74), and the cross-region correlation is 0.3829 with a 28.28 z-stat. There is little evidence of trading spreading from one region to the next.[8] The results are available from the authors upon request and are presented in a format similar to Table II (Panel A).

G. Telephone Trades as a Control Group

We previously exclude telephone trades from our study for two reasons. First, we want to concentrate on investors who are physically standing in the same

[8] In an earlier version of this paper, we performed yet another test to bolster our argument (not reported). We reran a correlated trading test, but ignored trades that happened before 11:00 a.m. each day. We found the same significant correlation patterns within a given region as we do at a daily frequency. Our test ignores the first hour and a half of the trading day. Thus, we eliminate the following scenario: investors go home each night, compare notes with investors who use other branch offices, return to trade in the morning, and place trades that reflect information obtained from the investors who use the other offices. As mentioned above, only tick-by-tick data could allow a true intraday study (and these data do not exist in the PRC). This said, we are now very confident that trades in our sample result from public information shocks.

room. Second, had we found a group-psychology effect, we planned to use the telephone trades as control group. Generally, an individual is either a telephone-trader or a branch-trader.

Although not reported, we repeat tests of correlated trading using only trades placed by telephone. There are few discernable differences between in-branch trading and telephone trades as far as correlated trading is concerned.

H. Possible Recommendations

A reader of this paper might hypothesize that brokers at the branch offices make recommendations and these recommendations cause investors to trade together. Such a hypothesis is extremely unlikely. First of all, all of the branch offices in our sample are part of the same brokerage company. Therefore, the same company would have to make different, and conflicting, recommendations in Guangdong and Shanghai. Second, the brokerage firm does not issue recommendations. Third, we have observed trading and found no evidence of recommendations. Fourth, the brokerage firm states that they do not issue recommendations. Finally, the staff on hand only advise investors on how to fill out order forms. They do not offer investment advice. Again, the firm and our own observations support our results.

I. Table III Revisited—at the Stock-Branch Level

An earlier version of this paper presented a regression similar to those regressions in Table III, except data was at the stock-branch level instead of the stock-region level. This disaggregated form runs into data density issues, but gives qualitatively similar results. The regression coefficient (z-stat) of the "near" investors was −0.0295 (−5.03). For the "far" investors, the values are 0.0042 (0.63). Clearly the coefficients are smaller—undoubtedly, due to noise inherent in the disaggregated form of the data. The sign of the regression coefficients remain the same. Statistical significance remains the same for the "near" investors but disappears for the "far" investors (also, undoubtedly due to data density issues).

J. Extreme Behavior

Some readers have suggested that the branch offices are not truly isolated. We believe they are, especially once we look at trading on a daily frequency (and after cutting out the trades before 11:00 a.m.).

We rule out (and have found no evidence of) special communication links between investor groups. Such links would have to consist of unheard of scenarios such as the following: one friend opens an account at one branch office, while his friend opens an account at another branch office. The two friends go every day to their respective branch offices armed with cellular phones. Each stands in the middle of his or her branch office, talks loudly on the phone, and announces to everyone what the other friend is observing at the other branch office. We find such explanations far-fetched.

K. Private Information Shocks

Some readers might think that private information shocks cause a group of investors to buy or sell a stock at the same time. It has even been suggested that our results could be the product of investment clubs whose members use (and meet at) the same brokerage office. If private information shocks were causing the observed trading, we would expect to see a positive correlation between net trades and returns in each branch office in our sample. In other words, as a group of investors demand liquidity, other market participants "agree" to provide it only at a higher price. We do not see this. We only see a positive correlation between net trades and returns at the "far" branches. We believe investors in the "far" branches are the ones who are least likely to have private information.

Also, we see that individuals at isolated offices (within the same region) buy and sell together. It is possible that, on occasion, private information could be impounded into stock prices by investors at two locations. But we see this happening time and time again. Remember, the correlation of net trades between two branch offices from the same region is positive and statistically so. Thus, our research design allows us to say, with a very high level of confidence, that there is a common, contemporaneous factor affecting the trading of isolated individuals. Calling this factor a "private information shock" simply does not fit the facts.

L. Our Model and Profitability

There are two ways to think about the rational expectations model reviewed in this paper. It is possible that nearby investors perceive they have more precise information about local companies. This perception could possibly come from overconfidence. If this were true, we would also see the same correlation of net trades and stock returns detailed in H_A and Table III. If the "overconfident" investors disregarded feedback signals from the market (i.e., they were not really better informed and did not actually outperform other investors), this correlation pattern could persist indefinitely. There would be no relationship between location and profitability.

If we believe the rational expectations model in the strictest sense, then we might believe that the better-informed investors should be better able to predict future returns. To test this hypothesis, we regress current net trades on future returns (i.e., current returns on lagged net trades). We use a generalized least squares methodology that allows for contemporaneous correlation between the stocks in our sample. Our results (not reported) show a positive correlation for Guangdong ("near") investors between net trades this week and returns next week. The results are not statistically significant. However, correlation of net trades and future returns for the Shanghai ("far") investors is significantly less than zero. This result is consistent with Coval and Moskowitz (2001).

M. Relevance of the Results

We end this section with a discussion of the relevance of our results. Some readers might believe that studying individual investors and/or PRC investors does not help to understand financial markets in the United States. We beg to differ.

As Hertz (1998) clearly describes, the PRC is a place where we, ex ante, expect to find group-psychology effects. If we fail to detect such effects in the PRC, it seems unlikely we would find them in the United States. That is, it is unlikely that individual and institutional investors in the United States are more prone to group-psychology effects than individuals in the PRC. More importantly, existing studies in the United States look at mutual funds. If institutions currently hold approximately 60% of the shares of the largest companies in the United States, then individuals hold approximately 40% of the shares. Suppose we divide investors in the United States into two groups (institutions and individuals). As discussed earlier, studying the trading behavior of one group can tell us about the other group if we believe the groups are systematically different and trade "against" each other (i.e., mutual funds vs. individuals).

IV. Conclusion

This paper provides an in-depth look at investor behavior. We show that individual investors engage in correlated trading behavior. We provide a number of tests designed to give us insight about the cause of correlated trading. We do this through experimental design. That is, we isolate groups of investors who are physically in the same room at the time they place trades. We are able to rule out group psychology as the predominant force driving investment decisions. If group psychology were important, then we would expect the trades of one isolated group to be uncorrelated with trades of other isolated groups (i.e., there is extremely low probability that two isolated groups would choose to buy a particular stock at the same time or sell a particular stock at the same time). Instead, we see isolated groups of investors in one region of the country tend to buy and sell together. Investors in another region of the country tend to buy and sell together. When one region is buying, the other is selling.

We present the testable hypotheses from a rational expectations model. We assume that investors who live near the firm receive more precise information (than those who live far from the firm) about future dividends. This assumption gives rise to a negative correlation of net trades for "near" investors and returns. Our data are consistent with such a negative correlation. The model is also consistent with the observed (positive) correlation of net trades for "far" investors and returns.

We use a principal component analysis to estimate the fraction of trading that can be explained by location-based effects. We show that trades from one region of the PRC load heavily on the main factor. Trades from another region load heavily (but oppositely) on the same, common factor. These results support the idea that public (or market-wide) information is a major determinant of trading decisions. The decision to buy or sell depends on location (and we observe this as a positive or negative loading on the factor). Our conclusions remain unchanged regardless of which methodology we use (correlations, regressions, or principal component analysis). Finally, a number of alternative hypotheses are explored.

We show what has traditionally been called "herding" in the finance literature is consistent with trade between asymmetrically informed agents. Thus, this paper gives insight into the structure of information in domestic stock markets. It also helps us understand the determinants of investment decisions.

Appendix A

This section reviews a simple, rational expectations model.[9] Our goal is to show that the net trading in our data is consistent with the predictions of a rational expectations model. We present an extremely sparse model and use the assumption that investors who live near a company receive more precise information about that company than investors who live far away from the company. The model and this one assumption provided a number of testable hypotheses.

A. Set-up of Model

The model has three periods, $t = \{0, 1, 2\}$, and a continuum of investors who are uniformly distributed on the closed interval from zero to one. Each investor maximizes the utility of his or her wealth at the end of period-2:

$$U_i(\tilde{W}_{2,i}) = -\exp\left[-\frac{1}{\gamma_i}\tilde{W}_{2,i}\right],$$

where γ_i is investor i's risk tolerance. For now, we assume all investors have the same risk tolerance, γ. The timing of the model is as follows. In period-0, investors receive an endowment of a riskless bond, $e_{riskless,i}$, and a risky asset, $e_{risky,i}$,. The riskless bond pays one unit of consumption in the final period. Without loss of generality, we assume the riskless interest rate is zero and the end of period price of the riskless bond is one: $p_{t,riskless} \equiv 1$ for $t = \{0, 1, 2\}$. The risky asset pays an amount $\tilde{u}$ in the final period, where $\tilde{u} \sim N(\bar{u}, 1/h)$. In order to avoid a fully revealing equilibrium, the aggregate endowment of the risky asset is not known to investors: $\tilde{e}_{risky} \equiv \int_i e_{risky,i} di$ and $\tilde{e}_{risky} \sim N(0, 1/\theta)$.

[9] The model in this paper most closely follows Kim and Verrecchia (1991), although the assumption about information asymmetry is most similar to that in Brennan and Cao (1997). For those interested in similar work, see Admati (1985) and Dvorak (2001).

In period-0, investors receive two signals about the future dividend of the risky asset, a public signal $\tilde{y}_{t=0} = \tilde{u} + \tilde{\eta}_{t=0}$ and $\tilde{\eta}_{t=0} \sim N(0, 1/\eta_0)$ and a private signal $\tilde{z}_{t=0,i} = \tilde{u} + \tilde{\varepsilon}_i$ and $\tilde{\varepsilon}_i \sim N(0, 1/s_{t=0,i})$. It is assumed that the precision of the private signal, $s_{t=0,i}$, is positive and bounded. After receiving their endowments and signals, investors trade so the price at the end of period-0 is consistent with their beliefs. In period-1, investors receive only a public signal about the future dividend of the risky asset, $\tilde{y}_{t=1} = \tilde{u} + \tilde{\eta}_{t=1}$ and $\tilde{\eta}_{t=1} \sim N(0, 1/\eta_1)$. After receiving the signal, investors trade. In the final period, the dividend of the risky asset is revealed: $\tilde{p}_{risky,t=2} = \tilde{u}$.

At the end of the first and second time periods (when trading takes place), investors solve for the optimal demand of the risky asset ($x_{riskless,t,i}$ and $x_{risky,t,i}$ for $t = \{0, 1\}$), given the information available to them at the time. We stack the endowments, prices, and holdings of the assets for notational simplicity:

$$\begin{aligned} E_i &\equiv [e_{riskless,i} \quad e_{risky,i}]' \\ P_t &\equiv [p_{riskless,t} \quad p_{risky,t}]' \\ X_{t,i} &\equiv [x_{riskless,t,i} \quad x_{risky,t,i}]'. \end{aligned}$$

The final period wealth of an individual investor can be written

$$\tilde{W}_{2,i} = P_0' E_i + (\tilde{P}_1 - \tilde{P}_0)' X_{0,i} + (\tilde{P}_2 - \tilde{P}_1)' X_{1,i}. \tag{A1}$$

B. Location Information Assumption—from Brennan and Cao (1997)

The precision of the private signal ($s_{t=0,i}$) is assumed to vary across investors. We further assume that investors who live near the headquarters of a traded stock company receive more precise private signals (on average) than those investors who live far away from the headquarters. The motivation behind this assumption can be thought of as follows: investors who live near the headquarters may have friends who work for the company, there may be more news stories about the company in the local paper, they may use the company's products more often, etc. In the United States, it is not hard to imagine that investors read more about their local telephone company than a telephone company across the country.

ASSUMPTION 1 (Location and private information; from Brennan and Cao 1997): *Investors can be divided into two groups, $i \in \{$"near", "far"$\}$, based on the distance they live from a company's headquarters. Investors who live near the headquarters of a firm receive (on average) more precise information (private signals in the model) than those who live far away.*

We can write the assumption more formally as

$$E[s_i \mid i \in near] > E[s_i \mid i \in far]. \tag{A2}$$

C. Solution to the Model

A brief description regarding the solution to the model is given here.[10] Investors make conjectures about the relationship between prices and information in the economy. In equilibrium, these conjectures are true. Prices of the risky asset are conjectured to be a linear function of the expected dividend (and all signals related to the dividend) as well as of the expected supply of the risky asset. Investors also know the precision of their own information (called $K_{i,t}$) at time t. The average precision in the market at time t is written simply K_t.

D. Equilibrium Holding Levels

The equilibrium holding of investor "i" in the risky asset at the end of period-0 is given by the expression below. Here s is the average precision of the private signal across investors. We take out the "t=" in the subscripts and leave only the time period for compactness:

$$\tilde{x}_{risky,0,i} = \gamma s_{0,i}\epsilon_i + \frac{\gamma(s_{0,i} - s)}{K_0}[h(\tilde{u} - \bar{u}) - \eta_0\tilde{\eta}_0 + \gamma s\theta\tilde{e}_{risky}] + \frac{K_{0,i}}{K_0}\tilde{e}_{risky}. \quad \text{(A3)}$$

And the equilibrium holding of investor "i" in the risky asset at the end of period-1 is given by the expression below.

$$\tilde{x}_{risky,1,i} = \gamma s_{0,i}\epsilon_i + \frac{\gamma(s_{0,i} - s)}{K_1}[h(\tilde{u} - \bar{u}) - \eta_0\tilde{\eta}_0 - \eta_1\tilde{\eta}_1 + \gamma s\theta\tilde{e}_{risky}] + \frac{K_{1,i}}{K_1}\tilde{e}_{risky}. \quad \text{(A4)}$$

Once the equilibrium holdings level at time $t = 1$ and $t = 0$ are found, the difference between these level is the net buying (selling) for investor "i" over the period. The absolute value of the net buying or selling gives us the volume for investor "i." If we aggregate individual volumes across all investors, we get an expression for total volume in the market. Note, the only event that causes the volume and net trading in this model is the release of a pubic signal in period-1.

E. Volume of Trade

The volume of total trade in the whole market caused by the public signal in period-1 is

$$Volume_{risky,t=0\to 1} \equiv \left(\frac{1}{2}\int \gamma\,|s_{t=0,i} - s|\,di\right)|\Delta\tilde{P}_{t=0\to 1}|. \quad \text{(A5)}$$

We see that volume and contemporaneous price movement (returns) are correlated. Volume is also correlated with dispersion of beliefs within the economy.

[10] For readers who are interested in a complete solution to the model and additional details, please see Kim and Verrecchia (1991).

This makes sense. If everyone has the same beliefs, then a release of public information can cause prices to move, but does not induce trading.

F. Net Trade

The change in holdings of the risky asset, for investor "i," caused by the public signal in period-1 is

$$Net_{risky,t=0\to 1,i} \equiv \Delta\tilde{x}_{risky,t=0\to 1,i}, \tag{A6}$$

$$= -\gamma(s_{t=0,i} - s)\Delta\tilde{P}_{t=0\to 1}. \tag{A7}$$

Again, s is the average precision of the private signal across investors and $s_{t=0,i}$ is the private signal of investor "i." We see that an investor updates his or her beliefs about the value of the risky asset only if the precision of his or her prior information is below (or above) the average precision in the economy. Investors with less precise information will tend to update more heavily than those with more precise information.

Appendix B

Table V (Panel A) shows the results of a principal component analysis. We see that the first principal component explains 31.83% of the variance of net trades, the second component explains 15.96% of the variance, etc. In order to estimate the significance of these percentages, we use a Monte-Carlo method.

Our data represent net trades and are in an 80×7 matrix (80 weeks and seven branch offices). If the net trades of each branch were uncorrelated with those of other branches (i.e., no common component), then each principal component should explain (1/7) or 14.29% of the total variance. However, even under the null of uncorrelated trading across branches, the first principal component may explain more than (1/7) of the total variance (when using our actual data). This is due to the fact that the percentage is bounded from below by (1/7) and biased upward.

We simulate our data by drawing from the original data with replacement. That is, we draw an 80×7 matrix of data where draws in the column-n come from the branch-n distribution of net trades. We then extract the first principal component and calculate the percent of variance explained. The empirical distribution of trades and sample size of 80 lead to a simulated first principal component that explains 20.65% of the variance. Our finding of 31.83% is significantly above this value.

We now filter out the first principal component and look at the significance of the second component. Since the first principal component from our random draws explains 20.65% on average, we expect the second component to explain (1/6) of the remaining variance of 79.35%. This fraction is 13.23%. Again, due to sample size, the percentage is actually bounded from below by (1/6) and is biased upward. Our second principal component is significantly higher than this number. We continue in a similar manner for the remainder of the components.

Alternative Methodology: We double-check our principal component results using a maximum likelihood factor analysis implemented by STATA. This software package offers two statistical tests. The first is a chi-squared test that there are three factors vs. no factors. This test has a 0.0000 p-value. The second test is a chi-squared test comparing three factors vs. more than three factors. This test has a 0.9134 p-value. We are reassured about the statistical significance presented in Table V (Panel A).

More importantly any rotation of principal components that STATA generates gives the factor-loading pattern that we report in Table V. That is, the Guangdong branches load one way on the first factor, while the Shanghai branches load with the opposite sign on the same factor. Thus, we are confident about our economic interpretation of a common, market-wide shock.

Appendix C

Feedback Trading and Stock Returns

This table presents a regression trading activity on past returns. Data represent stock (equity) trades placed by individual investors in the PRC between May 4, 1999 and December 4, 2000. This time period represents 80 trading weeks, or 388 trading days. Trades are placed at one of seven brokerage offices, and the office is responsible for maintaining the investors' account data. We concentrate on a sample of buys or sells of high-volume stocks that are listed and headquartered in Guangdong (listed on the Shenzhen Stock Exchange) and in Shanghai (listed on the Shanghai Stock Exchange). Coefficients are estimated by generalized least squares with an allowance for heteroskedasticity and cross-sectional correlation between branches. We form portfolios across branches within a given region. In Panel A, there are 3,950 observations for "near" trades and 3,950 observations for "far" trades. The number of observations equals 80 weeks times 50 stocks (twenty-five from each region), minus 50 observations due to the lag. In Panel B, there are 19,350 observations for each of the two groups. The number of observations equals 388 days times 50 stocks (25 from each region), minus 50 observations due to the lag. Examples of "near" trades are the net trades of Guangdong investors in Guangdong stocks. Examples of "far" trades are the net trades of Guangdong investors in Shanghai stocks.

	Investor Location	
	Near	Far
Panel A: Weekly Regressions		
Dependent variable	$Net_{i,t}/\sigma_{Net,i}$	$Net_{i,t}/\sigma_{Net,i}$
Independent variable	$r_{i,t-1}/\sigma_i$	$r_{i,t-1}/\sigma_i$
$\hat{\beta}$	−0.0082	0.0181
(z-stat)	(−0.52)	(1.14)
Panel B: Daily Regressions		
Dependent variable	$Net_{i,t}/\sigma_{Net,i}$	$Net_{i,t}/\sigma_{Net,i}$
Independent variable	$r_{i,t-1}/\sigma_i$	$r_{i,t-1}/\sigma_i$
$\hat{\beta}$	−0.0083	0.0125
(z-stat)	(−1.15)	(1.74)

REFERENCES

Admati, Anat R., 1985, A noisy rational expectations equilibrium for multi-asset securities markets, *Econometrica* 53, 629–657.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2002, Systematic noise, Working paper, University of California, Davis.

Bikhchandani, Sushil, and Sunil Sharma, 2000, Herd behavior in financial markets: A review, IMF working paper, March.

Brennan, Michael J., and H. Henry Cao, 1997, International portfolio investment flows, *Journal of Finance* 52, 1851–1880.

Choe, Hyuka, Bong-Chan Kho, and Rene M. Stulz, 1999, Do foreign investors destabilize stock markets? The Korean experience in 1997, *Journal of Financial Economics* 54, 227–264.

Coval, Joshua D., and Tobias J. Moskowitz, 2001, The geography of investment: Informed trading and asset prices, *Journal of Political Economy* 109, 811–841.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann, 1990, Positive feedback investment strategies and destabilizing rational speculation, *Journal of Finance* XLV, 379–395.

DeMarzo, Peter M., Ron Kaniel, and Ilan Kremer, 2002, Diversification as a public good: Community effects in portfolio choice, *Journal of Finance* 59, 1677–1715.

Devenow, Andrea and Ivo Welch, 1996, Rational herding in financial markets, *European Economic Review* 40, 603–615.

Dvorak, Tomas, 2001, Gross capital flows and asymmetric information, Working paper, Williams College.

Grinblatt, Mark, and Matti Keloharju, 2000, The investment behavior and performance of various investor types: A study of Finland's unique data set, *Journal of Financial Economics* 25, 43–67.

Grinblatt, Mark, and Matti Keloharju, 2001, How distance, language, and culture influence stockholdings and trades, *Journal of Finance* LVI, 1053–1073.

Grinblatt, Mark, Sheridan Titman, and Russ Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, *American Economic Review* 85, 1088–1105.

Hau, Harald, 2001, Location matters: An examination of trading profits, *Journal of Finance* LVI, 1959–1983.

Hertz, Ellen, 1998, *The Trading Crowd: An Ethnography of the Shanghai Stock Market* (Cambridge University Press, Cambridge).

Hirshleifer, David, and Siew Hong Teoh, 2003, Herd behavior and cascading in capital markets: A review and synthesis, *European Financial Management* 9, 25–66.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2002, Thy neighbor's portfolio: Word-of-mouth effects in the holdings and trades of money managers, Working paper, Stanford Business School.

Huberman, Gur, 2001, Familiarity breeds investment, *Review of Financial Studies* 14, 659–680.

Kim, Oliver, and Robert E. Verrecchia, 1991, Trading volume and price reactions to public announcements, *Journal of Accounting Research* 29, 302–321.

Kim, Woochan, and Shan-Jin Wei, 2002a, Foreign portfolio investors before and during a crisis, *Journal of International Economics* 56, 77–96.

Kim, Woochan, and Shan-Jin Wei, 2002b, Offshore investment funds: Monsters in emerging markets? *Journal of Development Economics* 68, 205–224.

Kumar, Alok, and Charles M. C. Lee, 2002, Individual investor sentiment and comovement in small stock returns, Working paper, Cornell University.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny, 1992, The impact of institutional trading on stock prices, *Journal of Financial Economics* 32, 23–43.

Lobao, Julio, and Ana Paula Serra, 2002, Herding behavior: Evidence from Portuguese mutual funds, Working paper, Instituto de Estudos Financeiros e Fiscais.

Nofsinger, John R., and Richard W. Sias, 1999, Herding and feedback trading by institutional and individual investors, *Journal of Finance* LIV, 2263–2295.

Oehler, Andreas, and George Goeth-Chi Chao, 2000, Institutional herding in bond markets, Working paper, Bamberg University.

Wermers, Russ, 1999, Mutual fund herding and the impact on stock prices, *Journal of Finance* LIV, 581–622.