THE JOURNAL OF FINANCE • VOL. LXIII, NO. 2 • APRIL 2008

Correlated Trading and Returns

DANIEL DORN, GUR HUBERMAN, and PAUL SENGMUELLER*

ABSTRACT

A German broker's clients place similar speculative trades and therefore tend to be on the same side of the market in a given stock during a given day, week, month, and quarter. Aggregate liquidity effects, short sale constraints, the systematic execution of limit orders (coordinated through price movements) or the correlated trading of other investors who pick off retail limit orders do not fully explain why retail investors trade similarly. Correlated market orders lead returns, presumably due to persistent speculative price pressure. Correlated limit orders also predict subsequent returns, consistent with executed limit orders being compensated for accommodating liquidity demands.

ACROSS ALL MARKET PARTICIPANTS, trades net to zero in each stock: There is a buyer for every seller. However, subgroups of investors may be net buyers or sellers in a given stock during a given period. The subgroup studied in this paper is a sample of more than 37,000 retail clients at one of the three largest German (and European) discount brokers—brokers that do not give investment advice. The clients' complete daily transaction records, available from February 1998 to May 2000, allow us to address three related questions. First, do retail investors trade more similarly than we would expect them to merely by chance? Second, what coordinates retail trades? Third, what role does correlated retail trading play in price formation?

We are not the first to consider these questions. Barber, Odean, and Zhu (2003) report that trades of clients at a U.S. discount broker are correlated at a monthly frequency and that past returns and trading volume help coordinate retail trades. Using the same data, Kumar and Lee (2006) report that retail trading imbalances help explain monthly stock return variation in addition to what can be explained with commonly used empirical asset pricing factors. Both studies, as well as contemporaneous work on the relation between retail trading and price formation (Kaniel, Saar, and Titman (2008), Andrade, Chang,

*Dorn is with the LeBow College of Business, Drexel University; Huberman is with Columbia Business School and CEPR. Sengmueller is with CentER—Tilburg University and University of Amsterdam. We thank Martin Bohl, Maurice Bun, Larry Glosten, Will Goetzmann, Charlie Himmelberg, Frank de Jong, Anne Dorn, Alexander Ljungqvist, Mark Seasholes, Rob Stambaugh (the editor), Suresh Sundaresan, Russ Wermers, an anonymous referee, seminar participants at the Helsinki School of Economics, and conference participants at the 2006 workshop on "The Architecture of Financial System Stability" in Capri (Italy), the 12th Mitsui Life Symposium on Financial Markets at the University of Michigan, and the 2006 WFA meetings in Keystone for their comments. We also thank Inessa Love for sharing her panel VAR code with us. Daniel Dorn acknowledges financial support from a Dean's research grant.

and Seasholes (2007), Barber, Odean, and Zhu (2006), and Hvidkjaer (2006)) essentially treat retail trades as homogenous.

What sets our study apart is that we examine different types of retail trades. We distinguish between speculative and other trades, and between market orders and executed limit orders. These distinctions are crucial; our answers to the questions outlined above vary depending on the types of trades considered. Moreover, since our focus is on the daily frequency, we are able to show not only that correlated speculative market trades are contemporaneously correlated with returns, but also that they lead returns.

Retail investors in our sample exhibit a stronger tendency to trade in tandem than the institutional investors studied in earlier papers (see Lakonishok, Shleifer, and Vishny (1992) (LSV), Wermers (1999), and Wylie (2005)). In a typical stock and quarter, 57% of the investors are on the same side of the market (when one would expect 50% of the funds to be on the same side of the market simply by chance). In contrast, LSV report that only 50.1% of the U.S. pension funds in their sample are on the same side of the market in a typical stock and quarter. Correlated trading among retail investors persists at the daily frequency, with 54.3% of the investors moving in the same direction.

The correlated execution of stale limit orders is a mechanical reason why retail investors end up on the same side of the market. For example, a price jump can trigger the execution of limit sell orders that were submitted days or even weeks apart. Indeed, we find that 56.2% of the limit order traders move in the same direction in a typical stock and day whereas only 53.9% of the market order traders do.[1] However, mechanical reasons such as limit order execution or short-sale constraints (see Wylie (2005)) explain only a fraction of the comovement observed in our sample.

We further distinguish market orders into orders deemed likely nonspeculative, that is, driven by savings, dissavings, or risk-sharing motives, and those deemed likely speculative, that is, driven by perceived information about the future stock price (similar to Barber and Odean (2002)). For example, the transaction records contain a variable that identifies the order as part of an automatic investment plan through which retail investors can gradually build or reduce positions in individual stocks and mutual funds at pre-determined dates (similar to ShareBuilder in the United States); such plan trades are likely driven by savings or dissavings motives and are thus classified as nonspeculative. Given that nonspeculative orders are coordinated either explicitly (through investment plan purchases) or implicitly (for example, through time clustering of paycheck distributions), it is not surprising to find that such orders are indeed correlated.

A more surprising result is that the observed comovement among speculative market order trades is 20% to 50% higher than that among nonspeculative market order trades. In other words, retail investors move together primarily

[1] Barber et al. (2003) proxy for limit order effects by filtering out buys on days with negative returns and sells on days with positive returns. They report that their results are unchanged for the reduced sample of transactions.

because they tend to place similar speculative bets, not because their nonspeculative trades are coordinated through, say, automatic investment plans.

What are the circumstances under which correlated trading is more pronounced? Other things equal, there is greater correlated trading in the sample in a given stock and day when more clients trade or when the market-wide trading volume for that stock-day is larger. Correlated trading also tends to be greater in larger stocks. In contrast, we find little evidence that the actions of sophisticated investors, or investors located close to one another or the firm, are responsible for the documented comovement.

Correlated decisions to trade individual stocks have implications not only for individual investor welfare (see, for example, Barber and Odean (2000)), but potentially for prices as well. The individual trades in our sample appear to aggregate to an economically meaningful total. On average, the absolute value of the order imbalances in a given stock and day—number of shares purchased minus number of shares sold by the sample investors—represents close to 1% of the market-wide trading volume. Moreover, the clients of this brokerage are probably representative of other individual investors, especially clients of other discount brokers.

Individual investors could play a role in price formation because their orders, taken together, demand liquidity from or supply liquidity to other market participants. Individual investors might move prices because their speculative trades reveal information about future prices (see Glosten and Milgrom (1985) and Kyle (1985)). And even if investors traded on signals unrelated to fundamental values as conjectured in models of noise trading and style investing, their trading could move prices as long as other investors are constrained from betting against them (see, for example, De Long et al. (1990a), Shleifer and Vishny (1997), or Barberis and Shleifer (2003)).

In our sample, days and weeks with heavy speculative market order buying are associated with high returns, while speculative market order selling is associated with low returns. For example, the tercile of stocks most aggressively bought by speculators outperforms the tercile of stocks most heavily sold by 1.7% during the day of portfolio formation. Based on speculative market orders, the sample investors are momentum traders buying stocks with high past returns and selling stocks with low past returns. Moreover, the observed speculative trading leads returns as the speculative buy portfolio continues to outperform the speculative sell portfolio after portfolio formation. Predictability does not appear to be due to return momentum; a panel vector auto regression (VAR) confirms that trade imbalances are persistent and lead returns.

A second VAR, for a subsample of trades that condition on the previous day's information, indicates that even same-day returns are driven by trade imbalances and are not due to reverse causation. Predictability could therefore be due to persistent speculative price pressure that is not anticipated by other market participants.

The distinction between market orders and executed limit orders turns out to be crucial for understanding the relation between individual investor trading and returns. In contrast to aggregated market orders, aggregated limit orders appear to be negative feedback trades. In large part, the negative correlations between limit order imbalances and past as well as contemporaneous returns are due to the strong mechanical relation between price movements and limit order execution (Linnainmaa (2003) elaborates on this point). Limit order imbalances are positively correlated with future returns, which is consistent with the hypothesis that limit order traders are compensated for accommodating temporary liquidity demands of other investors (see Kaniel et al. (2008)).

A failure to separate market and limit orders would lead us to classify retail investors as contrarian, similar to prior and contemporaneous papers studying individual investor behavior (see Grinblatt and Keloharju (2000), Nofsinger and Sias (1999), Griffin, Harris, and Topaloglu (2003), Barber et al. (2003), Kaniel et al. (2008), Goetzmann and Massa (2002), and Jackson (2003) for evidence on Finnish, United States, and Australian retail investors)—the mechanical correlation between limit order execution and returns masks the previously undocumented relations between speculative trading and returns.

The remainder of the paper proceeds as follows. Section I provides a description of the transaction records. Section II examines comovement among the sample investors. Section III relates correlated speculative trading to returns, and Section IV concludes.

I. Data

The analysis in this paper draws on complete daily transaction records for a sample of more than 37,000 clients at one of the three largest German discount brokers between February 1998 and May 2000. The client sample is drawn randomly from the entire active and former client population as of January 1999—hence, the sample is free of survivorship bias—and represents roughly half of the existing client population at that time. The broker is labeled as a discount broker because no investment advice is given. This is important because it rules out correlated trading due to broker recommendations.[2] If only by virtue of its size, the sample is likely representative of the broader population of German discount brokerage clients. At the time the sample was drawn, it represented 10% of the total number of accounts held at German discount brokers (Van Steenis and Ossig (2000)). In turn, discount brokerage clients constitute a substantial fraction of the German retail investor population. At the end of our sample period, there were 1.6 million clients at German discount brokers (Van Steenis and Ossig (2000)) relative to roughly 6.2 million German investors with any exposure to individual stocks at the end of 2000 (see Deutsches Aktieninstitut (2003)). Dorn and Huberman (2005) document that German discount brokerage clients are younger, better educated, more likely to be self-employed, and earn higher incomes than the typical retail investor, similar to the differences between U.S. retail investors who trade online and those who do not (see New York Stock Exchange (2001)).

[2] Initial public offerings (IPOs) are an exception. During the sample period, IPOs are prominently displayed on the broker's web site. Moreover, the broker acts as an underwriter in several IPOs, which could be interpreted as an endorsement even if no explicit recommendation is given.

The transaction records are complete in that they contain the opening position as well as all transactions of a client from the account opening date until the account closing date or May 31, 2000, whichever comes earlier. This allows us to reconstitute the entire portfolio of each client at the end of each day during the sample period. The typical transaction record consists of a unique identification number, an account number that identifies the client, a transaction date, a buy/sell indicator, order type (indicating whether an order is a limit or a market order), order channel (indicating whether the order was placed, for example, by phone, internet, or as part of an automatic savings plan), an exchange indicator (identifying the exchange on which the order is placed), the type of asset traded (common stock, for example), a security identification code, the number of shares traded, the gross transaction value, and the transaction fees.

In principle, brokerage clients can trade all the bonds, stocks, and options listed on German exchanges, as well as all the mutual funds registered in Germany by domestic and foreign issuers. To examine correlated trading, we focus on transactions in domestic stocks for which Datastream provides detailed pricing and volume information. At the end of the sample period, the typical client portfolio is worth DEM 45,000 (USD 23,000 at the average DEM/USD rate of approximately 2 DEM per USD during the sample period) and consists entirely of individual stocks and stock mutual funds. The Herfindahl–Hirschmann Index (HHI) of the typical equity portfolio is 13.5% at the end of the sample period, that is, clients hold the equivalent of seven stocks equally weighted.[3] In aggregate, the brokerage positions are worth DEM 4.6 billion, almost 90% of which is in equities. The aggregate trading volume, the sum of absolute values of purchases and sales during the sample period, is DEM 18.3 billion. Most of the trading occurs in individual stocks and options.

We use Datastream to retrieve daily returns adjusted for splits and dividends for all German exchange-listed stocks, all registered mutual funds, and the DAX 100 and the Neuer Markt 50 (Nemax 50) indices.[4] We also use Datastream to obtain market capitalization, market-to-book values, daily opening and closing prices, and aggregate trading volume (the number of shares traded across all German stock exchanges for each stock-day, Datastream equity code "VZ") for most of the German stocks in our sample. We identify German stocks in a two-step process. First, we focus on stocks that have a German country code (the first two digits of the stock's International Security Identification Number (ISIN) are "DE"). Second, we check the underlying company names and stock descriptions to eliminate foreign stock certificates whose ISIN starts with "DE". Company names and stock descriptions come from a list maintained by the Karlsruher

[3] The HHI is defined as the sum of squared portfolio weights. A portfolio consisting of n stocks equally weighted would have an HHI of $n \cdot (\frac{1}{n})^2$. Note that we assume stock mutual funds to consist of 100 equally weighted positions that do not overlap with other holdings of the investor. That is, the HHI of the portfolio of an investor holding one mutual fund is 1% and that of an investor splitting his money equally between two mutual funds is 0.5%.

[4] The Neuer Markt was the market segment for growth and technology stocks of the Frankfurt Stock Exchange. It was founded in March 1997 and closed 6 years later.

Kapitalmarktdatenbank (this database makes German Stock Exchange data available for academic purposes).

II. Correlated Trading

A. Baseline Results

To detect herding among U.S. pension funds, Lakonishok et al. (1992) (LSV) develop a herding measure that has been subsequently used to assess the correlated trading behavior of, for example, U.S. mutual funds (Grinblatt, Titman, and Wermers (1995) and Wermers (1999)), U.K. mutual funds (Wylie (2005)), different groups of Finnish investors (Kyrolainen and Perttunen (2002)), and a sample of U.S. discount brokerage clients (Barber et al. (2003)). The LSV measure, H_{LSV}, is defined as

$$H_{LSV} = \frac{1}{\sum_{t=1}^{T} N_t} \sum_t \sum_j \left(\left| br_{jt} - br_t^{LSV} \right| - E_t\left[\left| br_{jt} - br_t^{LSV} \right| \right] \right), \tag{1}$$

where N_t is the number of stocks traded during period t, the buyers ratio br_{jt} is the number of net buyers of stock j during period t divided by the number of active traders of j during t, and the period-average buyers ratio br_t^{LSV} is the number of net buyers aggregated across all stocks during t divided by the number of all active traders during t. Subtracting the period-average buyers ratio controls for aggregate shifts into or out of stocks, for instance, in response to liquidity shocks.[5]

Table I shows the baseline results for the sample of German brokerage clients, considering daily, weekly, monthly, and quarterly horizons during the entire sample period from February 1, 1998 to May 31, 2000. We only consider secondary market purchases of German stocks. Note that LSV measures are only calculated for stock-periods with at least two active traders.

There is correlated trading across all horizons and the correlation appears to increase with longer observation intervals. At a quarterly horizon, for example, the average LSV measure is 8.3% and the median LSV measure is 7%. Assuming that the average fraction of position changes that are increases is one half, 57% of the sample brokerage clients are on the same side of the market in a typical stock quarter (see Panel A, Column (1), of Table I). By contrast, LSV report that only 50.1% of their sample pension fund managers trade in the same direction in a typical stock-quarter. Persistent buying or selling over time is one explanation for the positive relation between the level of herding and the length of the observation interval; we return to this explanation later.

[5] To get a test statistic that is zero under the null hypothesis of no correlated trading, LSV subtract the expected value of the difference between the buyers ratio and the period-average buyers ratio,

$$E_t\left[\left| br_{jt} - br_t^{LSV} \right| \right] \equiv \sum_{k=0}^{I_{jt}} \left(\binom{I_{jt}}{k} \left(br_t^{LSV} \right)^k \left(1 - br_t^{LSV} \right)^{I_{jt}-k} \left| \frac{k}{I_{jt}} - br_t^{LSV} \right| \right).$$

Table I

The LSV Measure of Correlated Trading: Baseline Results

Daily, weekly, monthly, and quarterly LSV measures are calculated for the sample as in Lakonishok et al. (1992). The sample period is February 1, 1998 to May 31, 2000. To form the top trading quartile ("top quartile"), we take each stock and identify the quarter of trading periods with the highest number of active traders in the stock. The bottom three trading quartiles ("bottom quartiles") contain the remaining observations. The LSV statistics in Column (4) are calculated for trades in DAX 30 stocks—the stocks that make up the index of the 30 largest and most liquid German stocks. Cross-sectional LSV measures are calculated by averaging LSV measures first across the observations for a given stock, then across stocks. To calculate industry LSV measures, we assign to each stock the Level 6 Datastream industry classification (Level 6 is the most detailed classification). LSV measures are then calculated for each industry rather than for each stock. All estimates are significantly different from zero at conventional significance levels assuming that the observations are independent.

	(1) All observations	(2) Bottom trading quartiles	(3) Top trading quartile	(4) DAX 30 stocks	(5) Cross section	(6) Industry
Panel A: Quarterly						
LSV - mean	8.3%	7.5%	9.4%	12.6%	8.1%	6.9%
LSV - median	7.0%	6.6%	7.9%	11.3%	7.8%	6.3%
LSV - *SD*	15.4%	15.8%	13.5%	12.1%	9.2%	4.0%
Number of observations	3,288	1,992	870	212	639	88
Panel B: Monthly						
LSV - mean	6.4%	5.5%	8.7%	11.4%	6.2%	6.1%
LSV - median	5.0%	3.8%	6.9%	9.7%	6.0%	5.5%
LSV - *SD*	15.3%	15.5%	13.5%	11.4%	7.5%	3.6%
Number of observations	7,552	5, 348	1,995	584	631	87
Panel C: Weekly						
LSV - mean	5.4%	4.6%	7.5%	9.2%	6.0%	4.8%
LSV - median	4.3%	3.6%	5.6%	7.8%	5.2%	4.8%
LSV - *SD*	16.2%	16.9%	13.7%	13.5%	7.8%	2.7%
Number of observations	21,176	15,585	5,429	2,461	616	87
Panel D: Daily						
LSV - mean	4.8%	4.0%	7.1%	7.5%	6.1%	4.5%
LSV - median	4.3%	3.5%	5.3%	7.1%	4.8%	3.9%
LSV - *SD*	17.6%	18.5%	14.6%	16.3%	8.5%	4.8%
Number of observations	47,341	35,209	11,950	10,527	591	88

In contrast to studies of U.S. institutional investor trading such as LSV and Wermers (1999), the LSV measure is much higher when calculated across active trading periods. We define active trading periods stock by stock as the quartile of observations with the broadest participation by the sample investors as measured by the number of net traders in a given period. As a consequence, each stock is represented roughly equally in the top trading quartile. The average LSV measure is 9.4% across stock-quarters that belong to the top trading quartile (see Panel A, Columns (2) and (3)), that is, almost 60% of the traders are on one side of the market and 40% are on the other side of the market when one would expect a 50%–50% split. When measured over shorter intervals, the average LSV measure declines monotonically from 6.4% at a monthly horizon to 4.8% at a daily horizon (see Panels B-D, Column (1)). Again, the LSV measure is higher when calculated across active trading periods. On stock-days in the top trading activity quartile, 57.1% of the clients are on the same side of market as opposed to 54% in the bottom three quartiles (see Panel D, Columns (2) and (3)). The results are similar when the period-average buyers ratio is estimated as the ratio of DEM purchases to DEM trading volume, that is, using trading volume rather than traders. Given that trading activity is commonly thought to be a proxy for differences of opinion (see, for example, Harris and Raviv (1993)), the positive correlation between trading activity and correlated trading in our sample is remarkable.

Also in contrast to LSV and Wermers (1999), who report that correlated trading is higher in small stocks, our sample investors tandem trade into and out of large stocks. Column (4) of Table I shows that the LSV measures for trading in stocks contained in the DAX 30, the index of the 30 largest and most liquid German stocks, are at least 50% higher than the overall measures.

All reported LSV measures are highly statistically significant at conventional levels assuming, as LSV and subsequent papers do, that observations are independent across time and across stocks. These assumptions are not quite innocuous: If the investors' tendency to buy or sell is serially correlated, observations will not be independent across time. However, even if we assume that observations are only independent across stocks, the LSV measures continue to be strongly significant (see Column (5) of Table I). Another possibility is that the observed comovement across stocks is a by-product of the investors moving together into and out of industries. If so, comovement measured across industries should be stronger than across individual stocks. Column (6) of Table I reports the LSV measures calculated for Level 6 (most detailed) Datastream Industry Classifications; the measures for coarser Datastream classifications are similar. The industry LSV measures are all positive and significant, increasing as the observation period lengthens from days to quarters. Their magnitude, however, is lower than that for individual stocks.

B. Robustness Checks

We verify that the observed comovement is not due to instances of mechanically correlated trading such as secondary market purchases of new stock issues (the results are virtually unchanged when excluding the first 6 months of IPO trades), exchanges of one stock for another at the conclusion of a merger or an acquisition or the distribution of bonus shares (the database identifies such transactions separately and we omit them), or stock repurchases.

B.1. Short-Sale Constraints

The standard interpretation of the LSV statistics assumes that a sale of a stock is just as likely as its purchase. For this to be true, short sales must be allowed and routinely executed. In many contexts, however, investors are legally barred from selling short (for example, U.K. fund managers (Wylie (2005)) either promise not to sell short (for example, long-only institutional money managers) or simply follow long-only strategies (Barber et al. (2003)). Moreover, some stocks are hard to short, especially for retail investors. During our sample period, none of the major German retail brokers allows its clients to short stocks.[6]

Wylie (2005) reports that even under the null of no herding, the LSV measure becomes significantly positive when investors face short-sale constraints. Intuitively, if few of the sample investors hold shares in a given stock, random initial purchases of that stock by other sample investors might be classified as correlated purchases. To assess the potential bias induced by short-sale constraints, we adopt a procedure similar to that in Wylie (2005); the procedure is described in Appendix A.

Table II summarizes the simulation results. In the presence of short-sale constraints, LSV measures are upwardly biased. Even if trading were uncorrelated, one should expect to observe LSV measures of around 2% across all horizons, which is significantly different from zero (by comparison, Wylie (2005) finds a bias of 1.5% for a semiannual horizon). In other words, one would expect to see 52% of the sample investors on the same side of the market in a typical stock-period even if their trading were not coordinated other than by short-sale constraints.

In short, short-sale constraints help coordinate trading, but capture neither the full extent of the observed comovement nor the concentration of trading activity in relatively few stocks, especially at short horizons.[7]

B.2. Limit versus Market Orders

Could high LSV measures simply reflect the coordinated execution of stale or otherwise unrelated limit orders? During a high-return day, for example, the execution of limit sell orders will be coordinated by the rising stock price.

[6] The brokers point to the unlimited downside risk associated with short selling and the resulting threat of legal action by retail clients. According to paragraph 37d of the German Securities Trading Law (Wertpapierhandelsgesetz WpHG), a broker can be liable for damages arising from certain transactions if the broker failed to fully disclose the risks associated with the transaction—see the edition of http://www.faz.net, of May 24, 2002, "Leerverkäufe—leichter gesagt als getan" (Short-selling—easier said than done), last viewed August 15, 2002.

[7] Another way of assessing the effect of short-sale constraints is proposed by Wermers (1999) who computes a buy herding measure and a sell herding measure by conditioning the LSV calculation on stock-periods when the buyers ratio exceeds the period-average buyers ratio (for the buy measure) and on stock-periods when the buyers ratio is below the period-average buyers ratio (for the sell measure). As in Wermers (1999), we find stronger evidence for sell herding, which appears at odds with short-sale constraints driving the results.

Table II

Correlated Trading Due to Short-Sale Constraints

The bias of the LSV measure due to short-sale constraints is estimated similar to Wylie (2005). The original data set is resampled 1,000 times assuming that the sample investors trade randomly, but face short-sale constraints. Quartiles of observations based on trading activity are formed as in Table I. The mean and standard deviation of the LSV measures are calculated across the 1,000 simulations. All estimates are significantly different from zero at conventional significance levels.

	All observations	Bottom quartiles	Top quartile
	Panel A: Daily		
LSV bias - mean	2.31%	2.16%	2.75%
LSV bias - *SD*	0.06%	0.08%	0.08%
	Panel B: Weekly		
LSV bias - mean	1.89%	1.75%	2.28%
LSV bias - *SD*	0.08%	0.10%	0.11%
	Panel C: Monthly		
LSV bias - mean	1.75%	1.65%	2.01%
LSV bias - *SD*	0.10%	0.13%	0.15%
	Panel D: Quarterly		
LSV bias - mean	1.79%	1.60%	2.27%
LSV bias - *SD*	0.16%	0.20%	0.19%

Alternatively, the observed comovement could be a by-product of other traders picking off the sample investors' limit orders, that is, the sample investors might be providing liquidity to the market (see Kaniel et al. (2008)).

The data allow us to distinguish between executed market and limit orders; further, limit orders can be classified as regular, stop buy, and stop loss orders. Although there is no information about the limit level, the database identifies the limit expiration date as the day of order submission, the last trading day of the current month, the last trading day of the next month, or "good until cancelled." The majority of executed limit orders—55%—have a limit expiration date beyond the day of order submission. Presumably, the proportion of unmonitored limit orders is higher among orders with longer limit expiry dates.

To get a conservative estimate of how limit orders affect trading correlations, we calculate LSV measures separately for market orders and for limit orders (regular, stop-buy, and stop-loss). The results, reported in Table III, indicate that the LSV measures based on market orders are about 10% less than the LSV measures based on all orders. Moreover, LSV measures based on market orders are also smaller than the LSV measures based solely on limit orders; the difference between market order comovement and limit order comovement depends on the observation frequency.

At a quarterly frequency, there is little difference between market order-based LSV and limit order-based LSV. By contrast, at the daily frequency, the distortion of the LSV measure due to the correlated execution of limit orders is

Table III

Correlated Trading: Limit versus Market Orders

Daily, weekly, monthly, and quarterly LSV measures of correlated trading are calculated separately for market orders and for limit orders. The sample period is February 1, 1998 to May 31, 2000. Quartiles of observations based on trading activity are formed as in Table I. The LSV biases due to short-sale constraints, estimated by resampling the original data set 1,000 times (and assuming that the sample investors trade randomly, but face short-sale constraints), are very similar to those reported in Table II and thus omitted. All estimates are significantly different from zero at conventional significance levels assuming that the observations are independent across stocks and time.

	Market orders			Limit orders		
	All	Bottom trading quartiles	Top trading quartile	All	Bottom trading quartiles	Top trading quartile
			Panel A: Daily			
LSV - mean	4.4%	3.7%	6.2%	5.8%	4.9%	8.3%
LSV - median	3.9%	2.9%	4.8%	6.2%	5.8%	7.0%
LSV - *SD*	17.8%	18.8%	14.5%	18.4%	19.2%	15.8%
Number of observations	26,605	19,713	6,722	29,331	21,735	7,430
			Panel B: Weekly			
LSV - mean	5.1%	4.4%	6.9%	5.9%	5.1%	8.2%
LSV - median	4.4%	3.6%	5.2%	5.1%	4.3%	6.5%
LSV - *SD*	16.4%	17.2%	13.4%	17.3%	17.9%	14.9%
Number of observations	13,495	9,840	3,455	16,556	12,130	4,258
			Panel C: Monthly			
LSV - mean	5.9%	5.3%	7.3%	6.4%	5.6%	8.6%
LSV - median	4.8%	3.7%	6.3%	5.0%	4.2%	6.7%
LSV - *SD*	15.3%	15.8%	12.9%	16.0%	16.4%	14.1%
Number of observations	5,230	3,646	1,375	6,556	4,616	1,737
			Panel D: Quarterly			
LSV - mean	7.6%	7.6%	7.5%	7.5%	6.8%	9.0%
LSV - median	6.9%	6.9%	5.7%	6.1%	5.8%	7.0%
LSV - *SD*	16.0%	16.2%	13.3%	15.6%	15.7%	13.9%
Number of observations	2,328	1,302	575	2,955	1,769	767

considerable—during a given quarter, stock prices both rise and fall, triggering the execution of limit buy and limit sell orders, whereas on a given day, stock prices tend to move in one direction (if only because of a relatively small number of price fixings), triggering the execution of either limit buy or limit sell orders. On stock-days belonging to the top trading quartile, for example, the LSV measure based on limit orders averages 8.3% vs. 6.2% based on market orders.

B.3. Speculative versus Nonspeculative Orders

People trade for different nonspeculative reasons. For example, they might purchase stocks to save for retirement and sell stocks to make the down payment for a house, to diversify risk, to rebalance their portfolio, or to lower their tax bill (trading for nonspeculative motives does not imply that the trader lacks a view on the security he sells or buys). One would expect such nonspeculative trades in our data to be coordinated for several reasons.

In general, as employees tend to get their paychecks either around the turn of the month or the middle of the month, one might observe correlated purchases around paydays. In particular, the sample investors have access to automatic savings plans that allow them to gradually build or reduce positions in dozens of individual stocks and mutual funds at four pre-determined dates per month, similar to ShareBuilder in the United States.

Stock returns might also help coordinate nonspeculative trades. For example, after a price run-up, investors may choose to rebalance their position. Note, however, that tax motives are unlikely to trigger correlated selling of losers towards the end of the year since capital gains are essentially untaxed in Germany.[8]

Do the sample investors trade in tandem mainly due to overlapping nonspeculative trading motives or because they place similar speculative bets? To shed light on this question, we identify trades that are likely undertaken for nonspeculative reasons (such as savings, liquidity, or rebalancing) and compute LSV measures separately for these trades and trades deemed to be speculative.

We base our definition of what constitutes speculative and nonspeculative trading on the definition proposed by Barber and Odean (2002). They define speculative trades as "all profitable sales of complete positions that are followed by a purchase within three weeks and all purchases made within three weeks of a speculative sale." (p. 475) They require that complete positions be sold to filter out rebalancing sales, profitable positions be sold to filter out tax-loss sales, and that sales be closely followed by purchases to filter out liquidity sales. In other words, only rapid successions of sales and purchases are classified as speculative. This definition has an intuitive interpretation: An investor sells stock A and buys stock B because he believes that stock B will outperform stock A. We modify this definition in two ways. First, we include stock sales for a loss as well since, as indicated above, such sales are unlikely to be tax-motivated. Second, we classify trades in automatic savings plans as nonspeculative. Note that the definition is exhaustive, that is, each trade is classified as either speculative or nonspeculative.

[8] Capital gains in Germany are not taxable unless they are realized within a certain period known as the speculation period (6 months before 1999, 12 months from 1999 on). Unlike in the United States, however, German banks and brokers are not required to report client transactions to the German IRS during the sample period, making it impossible to adequately enforce this tax law. In 2004, the German Supreme Court (Bundesverfassungsgericht) ruled the speculation tax—dubbed the "dunce tax" in honor of the few investors who faithfully report speculation gains—unconstitutional during certain periods, upholding an earlier judgment by the German Finance Court (Bundesfinanzhof).

Table IV

Correlated Trading: Speculative versus Nonspeculative Orders

Daily, weekly, monthly, and quarterly LSV measures of correlated trading are calculated separately for speculative market orders and for nonspeculative market orders. Orders are classified as speculative or nonspeculative according to the definition of speculative trading described in detail in Section II.B.3. The sample period is February 1, 1998 to May 31, 2000. All estimates are significantly different from zero at conventional significance levels.

	(1) Nonspeculative market orders	(2) Speculative market orders
Panel A: Daily		
LSV - mean	4.7%	4.8%
LSV - median	5.8%	5.0%
LSV - *SD*	18.3%	18.5%
Number of observations	10,378	19,094
Panel B: Weekly		
LSV - mean	4.9%	6.1%
LSV - median	5.7%	6.0%
LSV - *SD*	17.7%	17.3%
Number of observations	7,629	10,703
Panel C: Monthly		
LSV - mean	4.8%	8.2%
LSV - median	4.5%	7.6%
LSV - *SD*	16.0%	16.6%
Number of observations	3,842	4,463
Panel D: Quarterly		
LSV - mean	5.5%	10.1%
LSV - median	4.8%	9.3%
LSV - *SD*	15.4%	16.6%
Number of observations	1,930	2,117

According to this definition of speculative and nonspeculative trading, roughly two out of five purchases and sales are classified as nonspeculative. Nonspeculative purchases represent little more than 30% of total purchase volume in DEM; this is partly due to savings plan transactions that are relatively small.[9]

Table IV reports the LSV measures, calculated separately for speculative and nonspeculative market orders. As anticipated, nonspeculative trades are significantly correlated at all horizons (see Column (1) of Table IV).

[9] In unreported calculations, we extend the definition of speculative and nonspeculative by using trades in assets other than stocks to help classify stock trades. For example, we classify the sale of an individual stock as diversification-driven if it is followed by the purchase of an equity mutual fund. This results in a higher estimate of how much trading is nonspeculative (still less than 50% of all trades and 40% of the purchase volume), but the estimates of correlated trading reported below are almost identical under the extended definition so we omit them.

Remarkably, however, speculative market orders are even more strongly correlated than nonspeculative trades, especially at longer horizons (see Column (2) of Table IV). At a quarterly horizon, for example, the LSV measure calculated for speculative trades is almost twice that for nonspeculative trades. These results suggest that correlated trading among the sample investors is mainly driven by speculative trading, that is, trading in response to perceived signals about the future price of a stock.

C. Determinants of Correlated Trading

Why people trade in speculative markets is an important unresolved issue, especially if one notes that on its heels follow other questions: Who trades, under what circumstances does trading increase, which securities tend to attract more trading, etc. Tables I through IV establish that clients of the broker studied tend to be on the same side of the trade—not merely because of net liquidity flowing into or out of stocks, short-sale constraints, limit orders, or otherwise mechanically correlated orders. Next, we examine the circumstances under which their herding is more pronounced.

People trade in an attempt to profit from signals that they perceive as indications of future price changes. Signals can be formal announcements whose timing is often known in advance (earnings announcements, for example). At the other extreme are signals that trigger trading but leave barely a trace in the business section of the newspaper, save a mention that trading volume and usually price change were exceptional. For instance, Cutler, Poterba, and Summers (1989) note that many big moves of the U.S. stock markets were not accompanied by major news, and on many days on which major news did appear, the stock market's move was not exceptionally large; Mitchell and Mulherin (1994) report that variation in news announcements explains relatively little variation in trading volume on U.S. stock exchanges.

An attempt to explicitly identify all or even a substantial fraction of the signals that motivate trading seems futile. However, one can use turnover—number of shares traded divided by number of shares held—on a given stock-day as a proxy for signal strength and examine the relation between the broker clients' tendency to be on the same side of the trade and the strength of the signal. Prior research suggests that trading activity reflects formal announcements such as recommendations or earnings announcements. Womack (1996), for example, reports that trading volume for U.S. stocks on recommendation days is at least twice the normal trading volume, on average; Frazzini and Lamont (2006)) report that trading volume on an earnings announcement day for U.S. stocks is 50% higher than normal volume, on average (see also Kandel and Pearson (1995)). Moreover, trading activity proxies for other signals and is available for a large sample of stocks in Datastream. By contrast, I/B/E/S coverage for Germany is limited to a subset of firms in Datastream and a subset of analysts. Trading volume in a given stock-day can be market-wide or can be limited to the broker's customers. The former is indicative of signals that are noticed and acted upon by the general population of market participants, including institutions, whereas the latter appeals more to the narrower subset of market participants: the broker's clients and those who behave similarly to them.

Investors who live close to company headquarters or those who live close to each other may be on the same side of the market because they receive correlated signals or interpret signals similarly. For example, Feng and Seasholes (2004) report that Chinese investors who live close to company headquarters tend to be on the same side of the market and trade differently from distant investors.[10]

For our sample, we match the home zip codes of German-based investors with the zip codes of the stock headquarters. The average distance between a trader on a given stock-day and company headquarters is a measure of how close the traders are located to the firm (see Coval and Moskowitz (1999)). The zip codes of investors and firms are translated into geographic longitude (lon) and latitude (lat) by matching them against a list of zip codes and the corresponding geographic coordinates for 6,900 German municipalities.[11] Given the geographic coordinates, the distance between trader i and firm j can be estimated as

$$d_{i,j} = earth\ radius \cdot \mathrm{acos}(\sin(lat_j)\sin(lat_i) + \cos(lat_j)\cos(lat_i)\cos(lon_j - lon_i)). \tag{2}$$

The average pairwise distance between traders for a given stock-period is a measure of how close investors are located to each other. If one of the traders has a mailing address outside Germany, we set the distance between him and other traders to 1,000 km; the results reported below are not sensitive to a particular parameter choice. To make the distance measures comparable across stocks, we normalize them by average distance between company headquarters and the full client sample.

A lack of investor sophistication may also help explain comovement. Malmendier and Shanthikumar (2007) report that small and presumably unsophisticated traders naıvely react to analyst recommendations. For example, small traders tend to make purchases following buy recommendations without taking into account the incentives of analysts to tout a stock; the authors report that such correlated purchases are followed by negative abnormal returns. More generally, unsophisticated investors might trade on signals not realizing that the information associated with the signal is already impounded in prices. To proxy for investor sophistication, we compute four variables for each stock-day: the logarithm of the average size of the trade, the logarithm of the average portfolio value (including all stocks and mutual funds) of the traders at the end of the prior month, the average Herfindahl-Hirschmann index (HHI) of the traders' stock portfolios, and the average experience of the traders measured by the difference between the observation month and the account opening date.

[10] Note that Feng and Seasholes (2004) examine investors who trade at the same physical location. In contrast, our sample investors submit orders either over the phone or on the internet.

[11] This list can be downloaded from http://www.astrologix.de/download/, last viewed August 26, 2006.

Presumably, groups of traders can be characterized as more sophisticated the larger the trades they place, the larger and better diversified their portfolios, and the more experienced the traders are, on average.

To examine the relative importance of the different explanations for what drives comovement, we estimate a panel regression of LSV measures on a given stock-day on a wide range of stock and investor characteristics as well as a full set of daily time dummies. Stock characteristics include the brokerage- and market-wide turnover on the stock-day, lagged stock returns, the logarithm of the stock's market capitalization at the end of the previous month, the stock's market-to-book value at the end of the previous month, and four dummy variables that indicate whether the stock belongs to one of the four major German stock indices—DAX 30 (the index of the 30 largest and most liquid stocks), MDAX (a mid-cap index of 70 stocks that rank behind the DAX 30 components in terms of size and liquidity), the SDAX (a small-cap index of stocks that rank behind the MDAX components), and the NEMAX 50 (the 50 largest and most liquid stocks that list on the Neuer Markt). The investor characteristics are computed for those who trade on a given stock-day: the normalized average distance between the traders and the stock, the normalized average of the pairwise distances between the traders, the logarithm of the average size of the trade, the logarithm of the average portfolio value at the end of the previous month, the average HHI, the average experience of the traders, the fraction of male traders, and the fraction of orders submitted online.

The results—reported separately for LSV measures based on all market orders, nonspeculative market orders, and speculative market orders (columns (1)–(3) of Table V)—suggest that for stock-days on which the broker's clients participate more aggressively, more of them tend to be on the same side of the trade. This relation is more pronounced for speculative trades than for nonspeculative trades.[12] A one-standard deviation increase in the number of clients who trade a stock speculatively on a given day (normalized by the number of clients who hold the stock at the end of the previous trading day) is associated with a 25% increase in the LSV measure for speculative trades; a one-standard deviation increase in the fraction of clients who trade a stock non-speculatively is associated with less than a 20% increase in the LSV measure for nonspeculative trades.

More of the speculative traders also appear to be on the same side for stock-day occasions on which market-wide turnover in the stock is higher: A one-standard deviation increase in market trading activity is associated with a 6% increase in the LSV measure, other things equal. By contrast, market trading activity is unrelated to the nonspeculative LSV measures.

The strong relation between trading activity and tandem trading is not driven by stock-days with a small number of traders. In fact, when we restrict our attention to stock-days on which at least five sample investors trade, the above inferences become stronger (results are not reported).

[12] Note that we exclude savings-plan transactions from our analysis of nonspeculative trades.

Table V

Panel Regression of Correlated Trading

The dependent variables are the LSV measure based on all market orders (column (1)), the LSV measure based on nonspeculative market orders (column (2)), and the LSV measure based on speculative market orders (column (3)). The unit of observation is a stock-day. "Average distance traders - firm" is the average distance between the traders and the stock headquarters on a given stock-day. "Average distance trader i - trader j" is the average pairwise distance between all traders. "Average trade size" is the average absolute value of all purchases and sales underlying the calculation of the dependent variable for that stock-day. To compute the "Average portfolio value" and "Average HHI," we look at the portfolio of each trader at the end of the previous month. "Brokerage trading activity" is the number of clients trading that stock on a given day divided by the number of clients holding that stock at the end of the previous trading day. "Market trading activity" is the number of shares traded across all German stock exchanges divided by the number of shares outstanding. "Lagged positive return" is the stock return on the previous day if the return was positive, and zero otherwise; "lagged negative return" is the stock return on the previous day if the return was negative, and zero otherwise. The market capitalization and market-to-book value used during month t are values recorded by Datastream for the last day of month $t-1$. All coefficients are expressed in percent. The standard errors, in parentheses, are corrected for heteroskedasticity and dependence within same-stock observations (see, for example, Williams (2000)). "Number of clusters" refers to the number of different stocks in the sample. Note: ***/**/* indicate that the coefficient estimates are significantly different from zero at the 1%/5%/10% level.

Dependent variable	(1) LSV measure All trades	(2) LSV measure Speculative trades	(3) LSV measure Nonspeculative trades
Constant	−15.933***	8.560	9.356*
	(2.121)	(5.448)	(5.371)
Average distance traders - firm	−0.461	−0.733*	−0.067
	(0.305)	(0.397)	(0.629)
Average distance trader i - trader j	0.208	0.449	−0.385
	(0.296)	(0.398)	(0.638)
ln(Average trade size)	0.197	−0.474*	−0.107
	(0.206)	(0.250)	(0.357)
ln(Average portfolio value)	0.071	0.203	0.178
	(0.145)	(0.184)	(0.299)
Average HHI	0.384	0.832	0.258
	(0.964)	(1.030)	(1.525)
Average experience	0.430*	−0.021	0.940*
	(0.227)	(0.214)	(0.554)
Fraction of online orders	−0.747	−1.213*	0.785
	(0.609)	(0.680)	(0.981)
Fraction of male traders	−0.467	0.059	−1.554
	(0.630)	(0.741)	(1.008)
Brokerage trading activity	25.258***	27.878***	82.680**
	(3.781)	(4.654)	(38.939)
Market trading activity	8.394***	6.017*	8.909
	(2.868)	(3.259)	(8.517)
Lagged positive return	−4.246	−5.603	−8.382
	(2.725)	(3.419)	(7.302)
Lagged negative return	−14.784***	−15.508***	−7.057
	(5.322)	(5.913)	(12.962)
ln(Market capitalization)	0.334*	0.348**	0.996**
	(0.184)	(0.154)	(0.487)
Market-to-book value	0.000	−0.004	0.001
	(0.003)	(0.004)	(0.005)
Ancillary statistics			
Index component controls	Yes	Yes	Yes
Daily time dummies	Yes	Yes	Yes
Number of observations	24,719	18,379	7,068
Number of clusters	354	323	214
R^2	6.0%	8.0%	10.6%

Consistent with the univariate results, correlated trading is higher in larger stocks. Controlling for market capitalization, the stock index dummies, indicating whether a stock belongs to one of the four major German stock indices, are insignificant.

Variation in the characteristics of traders and their portfolios largely fails to explain the observed variation in overall LSV measures and nonspeculative LSV measures. In contrast, other things equal, the average distance between traders and company, the average trade size, and the fraction of online orders are negatively correlated with LSV measures for speculative trades.

One interpretation is that the proximity between traders and firm, the average trade size, and the fraction of online trading are all proxies for the likelihood that the traders are exposed to the same signals (which are interpreted similarly, as the correlation between trading activity and LSV measures suggests). For instance, nearby investors receive correlated signals about a firm, for example, through local news outlets. Further, phone-based investors and smaller investors may not enjoy the low cost access to diverse sources of information available to their online and larger counterparts; instead, they are likely to receive only the stronger signals that are broadcast via traditional news media, which may coordinate their trading.

Given that measures of trading activity such as turnover are commonly thought to be proxies for differences of opinion (see, for example, Harris and Raviv (1993)), the positive correlation between trading activity and correlated trading in our sample, that is, the LSV measure, is remarkable, suggesting that discount brokerage clients place correlated speculative bets in response to trading stimuli. Such stimuli are stronger on days with more trading and for larger stocks.

III. Speculative Trading and Returns

Retail investors tend to be on the same side of the market. Correlated trading cannot be fully explained by liquidity considerations, short-sale constraints, or limit orders. Rather, retail investors act similarly because they have similar trading ideas about specific stocks and take similar bets. Effectively, the sample investors, who are probably representative of a much larger section of the retail investor population, bet against the rest of the market. What role do these bets play in price formation?

A. Hypotheses

Under the null hypothesis of an efficient and liquid market, prices will quickly reflect any information conveyed by the correlated trading of the sample investors. In particular, one would expect not to find a systematic relation between today's sample trades and tomorrow's prices.

Alternative theories about the relation between correlated trading and returns essentially differ in their assumption about whether traders have some kind of private information or not.

Models of liquidity or noise trading start from the premise that tandem traders are uninformed. In response to correlated liquidity trading, inventories of broker-dealers move away from desired levels as they accommodate their clients' demands. In order to return to desired levels, the broker-dealers change the securities' prices. Moreover, to the extent that their clients' demands are temporary, inventories and bid/ask prices will revert to their earlier levels. For instance, if the clients exert buying pressure, the broker-dealers will sell to clients from their own inventories but will also raise both the bid and the ask prices to entice clients to sell more to them (by raising the bid price) and to cause fewer purchases (by raising the ask price). As the clients respond to the price changes, inventories will increase and bid and ask prices will come back down (see, for example, Stoll (1978), Amihud and Mendelson (1980), and Grossman and Miller (1988)). Chordia and Subrahmanyam (2004) note that, in the short-term, liquidity imbalances could also be positively correlated with future returns when the imbalances are autocorrelated. For instance, liquidity purchases during a given period are associated with more purchases and hence price pressure during the following period.

Correlated trading could also be a by-product of the sample investors' accommodating other investors' liquidity demands (though this would presumably occur via limit orders). In this case, imbalances should be negatively correlated with contemporaneous returns if retail investors are compensated for their supplying liquidity and positively correlated with future returns as the price pressure eases (see also Campbell, Grossman, and Wang (1993)). It is possible, of course, that the sample investors act as liquidity providers during certain periods and as liquidity demanders during other periods, making it harder to detect systematic trading patterns driven by liquidity motives.

Models of noise trading (see, for example, De Long et al. (1990a), Shleifer and Vishny (1997), Hong and Stein (1999), Barberis and Shleifer (2003)) offer similar predictions as inventory-based models of liquidity trading with the observed investors in the role of liquidity traders. For example, if noise traders engage in predictable positive feedback trading, their net trading activity will be positively correlated with contemporaneous or past returns and negatively correlated with future returns as the price moves back to fundamental value (see De Long et al. (1990b)).

It is also possible that at least some sample investors have access to information about the future price of a stock (see also Kyle (1985) or Glosten and Milgrom (1985))—for instance, employees might buy or sell company stock in response to internal rumors about the success or failure of an important management initiative. In this case, trade direction and contemporaneous as well as future returns should be positively correlated as the information is incorporated into prices—the correlation should increase with the extent of informed trading (at least if price impact functions are linear (see Huberman and Stanzl (2004))).

Theory offers little guidance as to the period over which future returns should be measured, that is, how long it takes for private information or shocks due to liquidity or noise trader demands to be absorbed by the market. Constrained by the size of our sample—February 1998 to May 2000—we focus on the daily horizon and consider the weekly horizon in robustness checks.

B. Empirical Evidence

To capture the directional importance of the observed trading activity, we compute for each stock-day the ratio of net number of shares traded in our sample—number of shares bought minus number of shares sold—divided by one-half the total number of shares traded across all German stock exchanges as reported by Datastream (Datastream double counts the number of shares traded). For simplicity, henceforth we refer to this ratio as the *order imbalance*. For example, if sample investor A bought 200 shares of a given stock, B bought 300 shares, C sold 100 shares, and Datastream reported 20,000 shares traded in aggregate (i.e., 10,000 shares bought and 10,000 shares sold), then the order imbalance would be $(200 + 300 - 100)/10{,}000 = 4\%$. Across all stock-days, the directional trading by the sample investors represents close to 1% of the total trading volume, on average.

We choose the order imbalance rather than the buyers ratio used in the LSV computation because it better captures the importance of the aggregated sample trading in price formation. Moreover, the buyers ratio—and other measures of excess demand proposed by Lakonishok et al. (1992) and used, for example, by Wermers (1999)—can only be computed on days with trading activity in the sample; for the time-series analyses presented below, we would have to arbitrarily fill in missing values or drop them. By contrast, the order imbalance is simply zero on stock-days without sample trading participation (but positive aggregate trading activity).

All major results reported below are qualitatively similar and often quantitatively stronger when using alternative proxies such as buyers ratios or the net number of shares traded in the sample divided by the total number of shares outstanding (rather than normalizing by the total number of shares traded) for the directional trading activity of discount brokerage clients.

B.1. Portfolio Results

Each day, we rank all the stocks traded that day according to their order imbalance based on all market orders, speculative market orders, or nonspeculative market orders. We then assign them to one of three equally sized portfolios: the sell portfolio (the portfolio of stocks with the lowest, and typically negative, buy pressure), the hold portfolio, and the buy portfolio.

Panel A(i) of Table VI reports the average order imbalance based on all market orders for the three portfolios on the formation day as well as the 5 trading days before and after portfolio formation. The daily sell, hold, and buy portfolios consist of 35 stocks, on average. The observed net sales of stocks in the sell portfolio represent 1.42% of the market trading volume, and the observed net

Table VI

Portfolios Based on Market Order Imbalances

At the end of each trading day, we sort all stocks with at least one transaction within the brokerage according to their market order imbalance (number of shares bought minus number of shares sold divided by total number of shares traded across all German stock exchanges) and assign them to one of three equally sized portfolios: the sell portfolio (PF1, the portfolio with the lowest, and typically negative, order imbalance), the hold portfolio (PF2), and the buy portfolio (PF3). In Panel A (all market orders), sorts are based on imbalances from market orders executed during the period February 1998 through May 2000. Panels B and C are based on speculative and nonspeculative market orders.

Part (i) of each panel reports the average order imbalance of the stocks in each of the three portfolios (PF1, PF2, PF3) and of a zero-cost portfolio (PF3-PF1), which is long in buy-portfolio stocks and short in sell-portfolio stocks. Part (ii) reports the corresponding average return (in percent), in excess of the equally weighted return across all German stocks for which Datastream provides total returns, measured from the close of the previous trading day to the close of the current trading day. All averages are reported for the day of portfolio formation (0) and the 5 days previous (−5 −1) and subsequent (1 5) to portfolio formation.

Panel A: All Market Orders

(i) Average Order Imbalance (in%)

	−5	−4	−3	−2	−1	0	1	2	3	4	5
PF1	−0.098	−0.086	−0.090	−0.095	−0.108	−1.421	−0.118	−0.111	−0.090	−0.098	−0.100
	(0.011)	(0.011)	(0.013)	(0.012)	(0.014)	(0.058)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)
PF2	−0.019	−0.020	−0.019	−0.018	−0.023	−0.018	−0.017	−0.023	−0.021	−0.022	−0.021
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
PF3	−0.043	−0.041	−0.036	−0.026	−0.002	0.833	−0.037	−0.055	−0.050	−0.054	−0.063
	(0.009)	(0.010)	(0.010)	(0.009)	(0.009)	(0.047)	(0.008)	(0.010)	(0.009)	(0.010)	(0.010)
PF3–PF1	0.055	0.045	0.054	0.070	0.106	2.253	0.081	0.056	0.040	0.044	0.037
	(0.012)	(0.014)	(0.013)	(0.013)	(0.015)	(0.096)	(0.014)	(0.014)	(0.012)	(0.014)	(0.011)

(ii) Average Returns (in%)

	−5	−4	−3	−2	−1	0	1	2	3	4	5
PF1	0.040	0.023	0.119	0.041	−0.191	−0.486	−0.209	−0.072	−0.006	−0.046	−0.031
	(0.045)	(0.044)	(0.049)	(0.063)	(0.049)	(0.046)	(0.045)	(0.043)	(0.040)	(0.047)	(0.039)
PF2	0.042	0.061	−0.007	0.126	0.107	0.142	−0.033	−0.082	−0.060	−0.100	−0.067
	(0.032)	(0.036)	(0.031)	(0.037)	(0.037)	(0.037)	(0.037)	(0.036)	(0.035)	(0.033)	(0.036)
PF3	0.094	0.113	0.089	0.262	0.625	0.893	0.280	0.093	−0.086	−0.093	−0.013
	(0.051)	(0.054)	(0.055)	(0.059)	(0.064)	(0.078)	(0.058)	(0.055)	(0.045)	(0.048)	(0.047)
PF3–PF1	0.054	0.091	−0.030	0.221	0.816	1.379	0.489	0.165	−0.080	−0.047	0.017
	(0.055)	(0.054)	(0.059)	(0.079)	(0.071)	(0.090)	(0.067)	(0.049)	(0.045)	(0.052)	(0.050)

(*continued*)

Table VI—*Continued*

Panel B: Speculative Market Orders

(i) Average Order Imbalance (in%)

	−5	−4	−3	−2	−1	0	1	2	3	4	5
PF1	−0.072	−0.060	−0.050	−0.059	−0.061	−1.332	−0.078	−0.070	−0.049	−0.068	−0.064
	(0.010)	(0.012)	(0.011)	(0.011)	(0.012)	(0.055)	(0.012)	(0.010)	(0.009)	(0.012)	(0.010)
PF2	−0.014	−0.016	−0.017	−0.016	−0.018	−0.021	−0.016	−0.016	−0.017	−0.023	−0.015
	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.002)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)
PF3	−0.016	−0.026	−0.028	−0.012	0.012	0.738	−0.022	−0.032	−0.029	−0.032	−0.045
	(0.008)	(0.009)	(0.009)	(0.009)	(0.010)	(0.051)	(0.009)	(0.010)	(0.009)	(0.010)	(0.009)
PF3–PF1	0.056	0.034	0.022	0.047	0.074	2.070	0.056	0.038	0.020	0.036	0.020
	(0.011)	(0.013)	(0.012)	(0.013)	(0.014)	(0.095)	(0.013)	(0.014)	(0.010)	(0.013)	(0.010)

(ii) Average Returns (in%)

	−5	−4	−3	−2	−1	0	1	2	3	4	5
PF1	0.085	0.003	0.097	0.086	−0.255	−0.554	−0.208	−0.077	0.020	−0.045	−0.033
	(0.056)	(0.048)	(0.056)	(0.135)	(0.059)	(0.049)	(0.045)	(0.048)	(0.050)	(0.053)	(0.047)
PF2	0.023	0.065	0.061	0.122	0.154	0.122	−0.069	−0.087	−0.102	−0.115	−0.056
	(0.040)	(0.039)	(0.039)	(0.044)	(0.044)	(0.049)	(0.041)	(0.040)	(0.039)	(0.037)	(0.042)
PF3	0.213	0.181	0.145	0.417	0.752	1.114	0.351	0.124	−0.082	−0.062	0.048
	(0.060)	(0.062)	(0.068)	(0.078)	(0.081)	(0.097)	(0.065)	(0.064)	(0.054)	(0.055)	(0.055)
PF3–PF1	0.128	0.178	0.049	0.331	1.007	1.668	0.559	0.201	−0.101	−0.017	0.081
	(0.069)	(0.060)	(0.066)	(0.146)	(0.095)	(0.111)	(0.068)	(0.057)	(0.054)	(0.059)	(0.064)

(*continued*)

Table VI—*Continued*

Panel C: Nonspeculative Market Orders

(i) Average Order Imbalance (in%)

	−5	−4	−3	−2	−1	0	1	2	3	4	5
PF1	−0.048	−0.035	−0.040	−0.041	−0.055	−0.799	−0.068	−0.045	−0.038	−0.046	−0.052
	(0.010)	(0.009)	(0.010)	(0.009)	(0.011)	(0.037)	(0.011)	(0.009)	(0.009)	(0.008)	(0.008)
PF2	−0.005	−0.005	−0.002	−0.004	−0.003	−0.002	−0.003	−0.004	−0.006	−0.006	−0.004
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
PF3	−0.005	−0.005	−0.005	0.002	0.002	0.499	−0.003	−0.008	−0.006	−0.014	−0.011
	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)	(0.025)	(0.005)	(0.006)	(0.008)	(0.007)	(0.006)
PF3–PF1	0.043	0.030	0.035	0.043	0.057	1.298	0.065	0.037	0.032	0.031	0.041
	(0.011)	(0.010)	(0.010)	(0.010)	(0.011)	(0.049)	(0.012)	(0.010)	(0.011)	(0.010)	(0.008)

(ii) Average Returns (in%)

	−5	−4	−3	−2	−1	0	1	2	3	4	5
PF1	0.150	0.231	0.237	0.226	0.138	0.014	−0.134	−0.020	−0.019	−0.056	−0.042
	(0.065)	(0.065)	(0.061)	(0.072)	(0.075)	(0.065)	(0.059)	(0.055)	(0.043)	(0.051)	(0.047)
PF2	0.123	0.153	0.125	0.162	0.341	0.398	0.038	−0.081	−0.070	−0.103	−0.081
	(0.046)	(0.052)	(0.046)	(0.056)	(0.064)	(0.064)	(0.046)	(0.045)	(0.043)	(0.049)	(0.042)
PF3	−0.002	0.032	0.062	0.187	0.447	0.471	0.147	0.025	−0.090	−0.102	−0.061
	(0.057)	(0.052)	(0.055)	(0.060)	(0.060)	(0.068)	(0.056)	(0.058)	(0.051)	(0.049)	(0.049)
PF3–PF1	−0.152	−0.199	−0.175	−0.039	0.309	0.457	0.281	0.046	−0.071	−0.046	−0.019
	(0.073)	(0.075)	(0.080)	(0.082)	(0.078)	(0.077)	(0.075)	(0.070)	(0.060)	(0.064)	(0.056)

purchases of stocks in the buy portfolio represent 0.83% of the market trading volume, on average.[13]

Net trades are persistent, mostly because the sample investors tend to sell stocks in the sell portfolio before and after the portfolio formation day. For example, the average order imbalance in the sell portfolio stocks on the day after portfolio formation is −0.12%, which is significantly less than zero.

Panel A(ii) of Table VI reports the equally weighted average return across all stocks in a given portfolio and day—measured from close to close—in excess of the equally weighted return across all stocks that day. We choose an equally weighted benchmark rather than a value-weighted benchmark because the brokerage portfolio is tilted towards smaller growth and technology stocks that exhibit much higher returns than German blue-chip stocks during the sample period. The results are qualitatively similar when we use stock-specific benchmarks such as momentum or industry portfolios and therefore these results are not reported.

Panel A(ii) of Table VI indicates that when using market orders, the sample investors buy stocks that appreciated relative to other stocks in the days leading to the purchase and sell stocks that depreciated in the days leading to the sale. Thus, the sample investors appear to be short-term momentum traders.

On the day of portfolio formation, the buy portfolio posts an average excess return of 0.89%; in contrast, the sell portfolio posts an average excess return of −0.49%. More interestingly, stocks bought by the sample investors on the portfolio formation day continue to do well and stocks sold continue to do poorly. For example, on the first day after portfolio formation, stocks in the buy portfolio post significant excess returns of 0.28%, and stocks in the sell portfolio post significant excess returns of −0.21%, on average. The buy minus sell portfolio, that is, the zero-cost portfolio that is long the stocks in the buy portfolio and short the stocks in the sell portfolio, has an economically and statistically significant return of 0.49% on the next trading day.

Panels B and C of Table VI show the corresponding results when the order imbalance is estimated separately for speculative and nonspeculative market orders. The results are qualitatively similar—both speculative and nonspeculative order imbalances are positively correlated with past returns and lead future returns—but speculative order imbalances appear to be more highly correlated with past, contemporaneous, and future returns than nonspeculative order imbalances.

[13] The average sale is substantially larger than the average purchase (as in Barber and Odean (2000)), which explains why the average order imbalance across all portfolios in Table VI is slightly negative. One explanation for why sales tend to be larger than purchases is that the sample investors sell some of their holdings acquired before the start of the sample period. Another explanation is that transfers of holdings between brokers—a sample investor who wants to switch to the sample broker will generally transfer his holdings to the new account rather than sell at the old broker and buy back at the new broker—are not treated as purchases. The number of stock transfers into the sample broker outnumbers the number of stock transfers out of the broker three to one; in addition, many IPO allocations occur before the start of secondary trading and are thus not considered as purchases either.

Results based on weekly portfolio sorts are similar and thus not reported. The documented correlations between the order imbalance and contemporaneous as well as future returns admit several explanations. First, the sample investors are net buyers of stocks whose returns exhibit short-term momentum. Second, and more generally, the sample investors trade on information about future prices (other than past returns) and are correct, on average. Third, the observed trades exert serially correlated price pressure: Higher net purchases today are associated with higher returns tomorrow because higher purchases today are also associated with higher purchases tomorrow that affect prices tomorrow.

B.2. Vector Auto Regression Results

To formally capture the dynamics of the trading-return relation and discriminate between the candidate explanations for the correlation between sample order imbalances and future returns, we run two-variable panel vector auto regressions (VAR) of order imbalances and returns at the individual stock level. First, we estimate a reduced-form panel VAR of the form

$$r_{i,t} = a^r + b_1^r r_{i,t-1} + b_2^r r_{i,t-2} + \cdots + c_1^r x_{i,t-1} + c_2^r x_{i,t-2} + \cdots + \alpha_i^r + \gamma_t^r + v_{i,t}^r \qquad (3)$$

$$x_{i,t} = a^x + b_1^x r_{i,t-1} + b_2^x r_{i,t-2} + \cdots + c_1^x x_{i,t-1} + c_2^x x_{i,t-2} + \cdots + \alpha_i^x + \gamma_t^x + v_{i,t}^x, \qquad (4)$$

where $r_{i,t}$ is the return of stock i measured from the closing price on day $t-1$ to the closing price on day t (prices are adjusted for splits and dividends), $x_{i,t}$ is the order imbalance in stock i on day t, α_i^r and α_i^x are return- and buy pressure-fixed effects for stock i, and γ_i^r and γ_i^x are aggregate shocks (time dummies). The system is estimated using GMM with the minimum set of instruments (see Gilchrist and Himmelberg (1998) and Love (2001)).

Table VII reports the coefficient estimates of a five-lag VAR system where the order imbalance variable is constructed from all market orders. The coefficient estimates confirm the inferences from the univariate portfolio results: The net trading activity of the sample investors positively responds to past returns, is persistent, and leads returns.

If short-term price momentum accounted for the relation between order imbalances and future returns, one would expect the relation to vanish once we condition on past returns. This is not the case.

To assess the magnitude of the dynamic relations, we compute impulse responses under the assumption that both contemporaneous and lagged buy pressure affects returns, but that only lagged returns affect buy pressure. Under this assumption, a positive one-standard deviation shock to order imbalances, which corresponds to sample net purchases that represent 1.1% of market-wide volume, results in a same-day excess return of 0.2% and a next-day excess return of 0.05%. The cumulative responses of the system to one-standard deviation innovations in returns and buy pressure are graphed in Figure 1, with the lower left-hand panel showing the responses of returns to

Table VII

Vector Auto Regressions of Market Order Imbalances and Returns

The dynamics of daily returns $r_{i,t}$ and order imbalances $x_{i,t}$ are captured in a two-equation panel VAR, as described in equations (3) and (4). Returns are defined as the percent raw returns measured close-to-close. The order imbalance is the fraction of market-order net buys of stock i as a percentage of all trading in that stock on day t as reported by Datastream.

		Coef	SE	t-stat
		Dependent variable: $r_{i,t}$		
Lags of $x_{i,t}$				
1		0.0575	0.0087	6.63
2		0.0168	0.0085	1.97
3		−0.0247	0.0094	−2.63
4		−0.0109	0.0093	−1.17
5		0.0265	0.0094	2.82
Lags of $r_{i,t}$				
1		0.0009	0.0050	0.17
2		−0.0208	0.0044	−4.78
3		−0.0186	0.0039	−4.76
4		−0.0023	0.0037	−0.62
5		−0.0037	0.0038	−0.98
		Dependent variable: $x_{i,t}$		
Lags of $x_{i,t}$				
1		0.0148	0.0051	2.88
2		0.0082	0.0052	1.55
3		0.0016	0.0047	0.34
4		0.0020	0.0044	0.45
5		0.0085	0.0047	1.81
Lags of $r_{i,t}$				
1		0.0056	0.0007	7.69
2		−0.0001	0.0007	−0.11
3		0.0014	0.0007	2.07
4		0.0014	0.0007	2.04
5		0.0010	0.0007	1.45
N	228,843			

order imbalance shocks. Figure 2 compares this response across order types, all market orders, as well as speculative and nonspeculative market orders. As suggested by the univariate portfolio results, stock returns are more responsive to speculative order imbalance shocks than to nonspeculative shocks (to conserve space, the coefficient estimates of the corresponding VARs are not reported).

These results are robust to different variable specifications such as using buyers ratios as proxies for directional trading activity. To make the stock-level VARs more comparable to the portfolio results, we also estimate a Hasbrouck (1991)-style VAR by defining an order imbalance variable that is +1 on a given stock-day if the stock is part of the buy portfolio, −1 if the stock is part of the sell portfolio, and 0 if the stock is part of the hold portfolio or not traded that day.

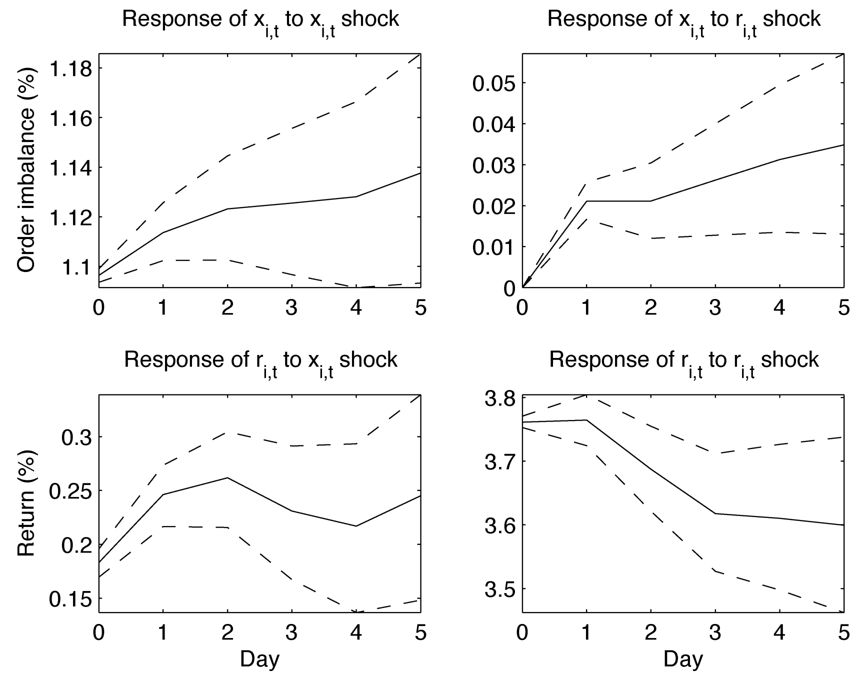

Figure 1. Impulse responses: market orders. The graphs show the cumulative responses of returns $r_{i,t}$ and order imbalances $x_{i,t}$ (constructed from market orders) to one-standard deviation shocks, based on the coefficient estimates in Table VII. The assumed ordering is (x, r), that is, shocks to returns are assumed to have no contemporaneous effect on order imbalances. Dashed lines indicate 95% confidence interval from Monte Carlo with 1,000 repetitions.

In this VAR system, a shock corresponding to a move from the hold to the buy portfolio is associated with a same-day excess return of almost 1% and a next-day excess return of more than 0.3% (detailed results are not reported), assuming, as above, that order imbalance shocks affect same-day returns but that return shocks do not affect same-day trading.

The assumption that returns on day t do not affect order imbalances on day t is strong in that it rules out intraday momentum trading or a common shock driving both returns and trading. Next, we examine trading in an institutional setting in which this assumption likely holds, which allows us to attach a causal interpretation to the VAR results.

Throughout the sample period, the broker recommends that clients submit their floor orders (orders routed to the floor of the Frankfurt Stock Exchange or one of the regional exchanges) between 10 a.m. and 11 a.m. at the latest on a given day to make sure that their orders be considered for same-day execution. This recommendation is motivated by the fact that, except for the largest stocks that are continuously traded in the electronic limit order book Xetra, there may only be a small number of price fixings on a given day. At the beginning of 1999, the broker introduced a trading platform that extended

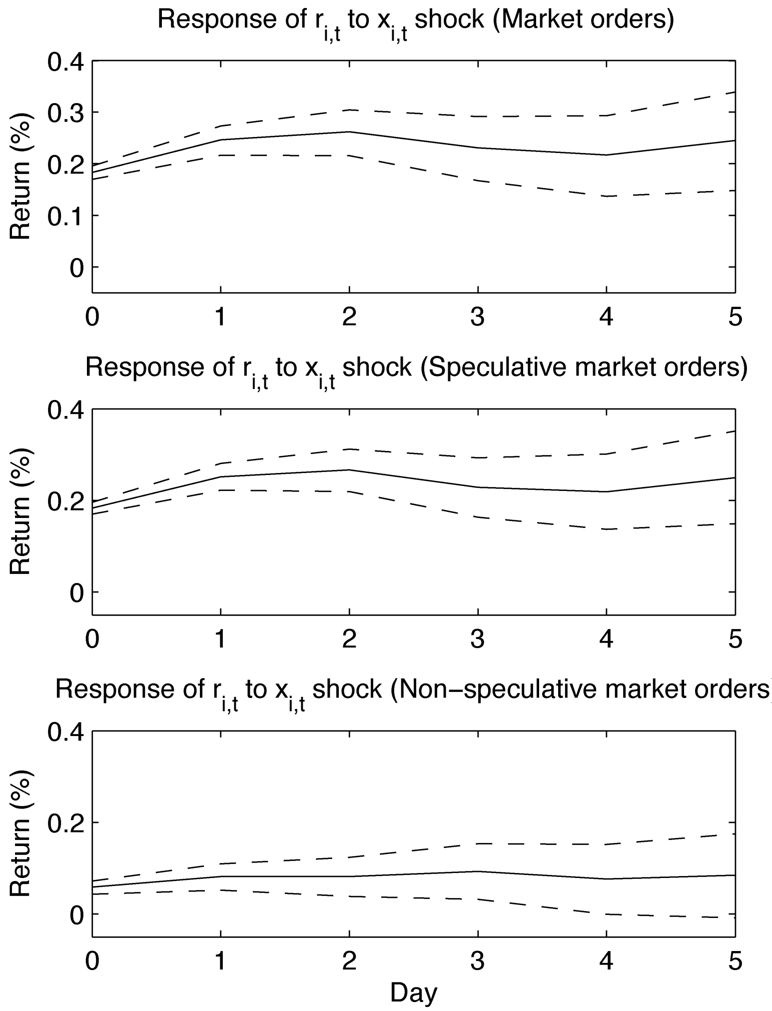

Figure 2. Comparing responses of $r_{i,t}$ to $x_{i,t}$ shock. The graphs compare the cumulative response of returns to a one-standard deviation in market-order imbalances and two subgroups thereof, speculative and nonspeculative order imbalances. All graphs assume the ordering (x, r), that is, shocks to returns are assumed to have no contemporaneous effect on order imbalances. Dashed lines indicate 95% confidence interval from Monte Carlo with 1,000 repetitions.

quasi-continuous trading to a much larger universe of stocks. Through this continuous trading platform, a brokerage client submits an order, the market maker immediately replies with a quote, and the client can decide whether to place the order or not. An order channel variable allows us to distinguish floor orders from those sent to the continuous trading platform. Between January 1999 and May 2000, orders submitted to the continuous trading platform account for 37% of all market orders. Here, we focus on the 63% of the market orders routed to the floor of the Frankfurt Stock Exchange or one of the regional exchanges.

Partly because it is difficult to condition floor trades on intraday returns, and partly because the continuous trading platform makes it easy to react to intraday returns or news releases from 1999 onwards, it is reasonable to assume that the sample floor trades from 1999 onwards are not submitted in response to intraday returns or news releases. In other words, the assumption that returns on day t do not affect order imbalances on day t is likely to hold when order imbalances are estimated for floor trades from 1999 onwards only.

We therefore estimate the panel VAR

$$r_{i,t}^{OC} = a^r + b_1^r r_{i,t-1}^{OC} + b_2^r r_{i,t-2}^{OC} + \cdots + c_1^r x_{i,t-1}^F + c_2^r x_{i,t-2}^F + \cdots + \alpha_i^r + \gamma_t^r + v_{i,t}^r \quad (5)$$

$$x_{i,t}^F = a^x + b_1^x r_{i,t-1}^{OC} + b_2^x r_{i,t-2}^{OC} + \cdots + c_1^x x_{i,t-1}^F + c_2^x x_{i,t-2}^F + \cdots + \alpha_i^x + \gamma_t^x + v_{i,t}^x, \quad (6)$$

where ${r^{OC}}_{i,t}$ is the return of stock i measured from the opening price on day t to the closing price on day t, and ${x^F}_{i,t}$ is the order imbalance in stock i on day t calculated on the basis of market orders submitted to the floor of the Frankfurt Stock Exchange or one of the regional exchanges from January 1999 to May 2000 only (i.e., excluding market orders executed on the continuous trading platform). We use open-to-close returns instead of close-to-close returns to address the possibility that floor trades executed on day t reflect news releases that become public between the close on day $t-1$ and the open on day t.

Figure 3 shows the cumulative impulse responses for this system. A positive one-standard deviation shock to order imbalances based on floor market orders (lower left-hand panel) is associated with a statistically significant same-day excess return of more than 0.1%. After 5 trading days, however, the cumulative return effect of the order imbalance shock is no longer significant at the 5% level.

These results suggest that the correlation between the sample trading activity and returns is due, above all, to price pressure. Given that the net trading activity in the sample is serially correlated, it appears that order imbalances predict future returns because net purchases today are associated with net purchases tomorrow, which in turn exert price pressure. Given the size of our sample and the noisiness of returns, it is difficult to detect longer-term reversals and hence to tell whether the price impact is temporary or permanent. Though our data do not allow us to definitely rule out informed trading as an explanation for the observed correlation, the previous analysis of the determinants of correlated trading gives us little reason to suspect that informed trading plays an important role. Moreover, we find in unreported results that clients who trade more speculatively fail to outperform their peers.

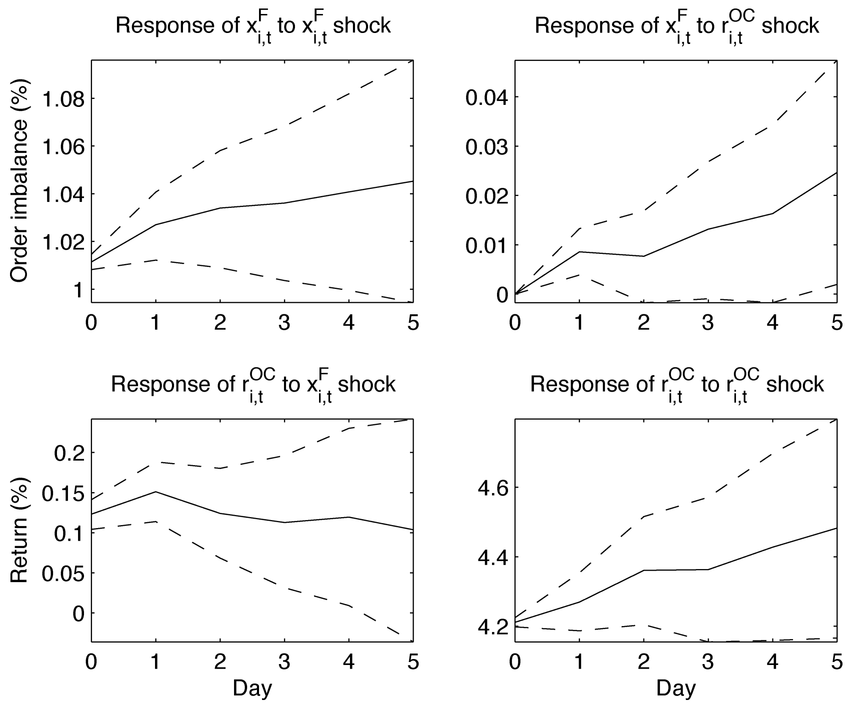

Figure 3. Impulse responses: floor trades. The graphs show the cumulative responses of open-close returns ${r^{OC}}_{i,t}$ and order imbalances ${x^F}_{i,t}$ (constructed from floor market orders) to one-standard deviation shocks. The impulse responses are based upon the estimates (not shown) of the VAR in equations (5) and (6). The assumed ordering is (x^F, r^{OC}), that is, shocks to returns are assumed to have no contemporaneous effect on order imbalances. Dashed lines indicate 95% confidence interval from Monte Carlo with 1,000 repetitions.

C. Discussion

At first glance, the results in the previous section appear to be at odds with the contrarian behavior reported for retail investors in Finnish stocks (Grinblatt and Keloharju (2000)), U.S. stocks (Nofsinger and Sias (1999), Griffin et al. (2003), Barber et al. (2003), Kaniel et al. (2008)), the S&P 500 index (Goetzmann and Massa (2002)), and Australian stocks (Jackson (2003)).

While differences in the data might account for some discrepancies, we suspect that the explanation lies in our focus on market orders in general, and speculative market orders in particular—all the papers referenced above combine market and limit orders. The advantage of focusing on speculative market orders is that they reflect investors' unadulterated opinion about future prices in a timely manner.

Limit orders may tell a different story, especially unmonitored limit orders, which provide liquidity in response to large price movements in an almost automatic fashion. To illustrate, let us start from a given set of limit buy and sell orders at the beginning of day zero and assume that these orders are unmonitored, that is, they will remain unchanged in the book unless they are executed. A large price decrease on day zero will make limit buy orders execute on day zero, implying a mechanical negative correlation between limit trading and contemporaneous returns. On day 1, unless the price increases by more than it decreased the previous day, only limit buy orders will be executed, and only if prices keep going down. Thus, unmonitored limit orders can explain the negative correlation between limit order imbalances and past returns. Without data on limit order submission, it is difficult to distinguish between active and mechanical contrarian behavior (Linnainmaa (2003) elaborates on this point).

Thus, one would expect the relation between limit orders and returns to be substantially different from the relation between market orders and returns. Table VIII confirms this conjecture.[14] The same-day return of a zero-cost portfolio long in stocks bought through limit orders and short in stocks sold through limit orders during that day is negative and highly significant (−1.45%). There is some evidence for a price reversal on subsequent days: On days 2 and 3 after formation, the zero-cost portfolio posts positive returns of 0.14% and 0.08%. If we combine all trades executed through the brokerage (not shown), the negative same-day effect of limit orders dominates: The mechanical negative correlation between limit orders and returns overwhelms the positive correlation between speculative market orders and returns.

Considering only limit orders, it is tempting to conclude that individuals follow a contrarian strategy at a daily and weekly frequency. The zero-cost portfolio based on limit orders has significantly negative returns on all 5 days prior to portfolio formation. It seems likely that the classification of retail investors as contrarians in the prior literature is due to the investors' use of limit orders. Without data on both submitted and executed limit orders, however, the negative correlation between imbalances and past returns is hard to interpret as it could be a reflection of "true" negative feedback trading or an artifact of unmonitored limit orders.

The reported correlations between limit order imbalances and returns suggest that retail limit traders do indeed provide liquidity to other market participants. However, the strong and possibly mechanical effect of limit orders masks the correlation between speculative trading and returns. Such heterogeneous trading strategies might help explain conflicting evidence on imbalances and returns (see, for example, Kaniel et al. (2008) vs. Andrade et al. (2007) or Linnainmaa (2003)).

[14] Although the market as a whole goes up during our sample period, stock-days with negative returns outnumber those with positive returns (in absolute terms, however, positive returns are larger than negative returns). This explains the positive average order imbalance across all limit-order portfolios in Table VIII. The greater persistence of order imbalances based on limit orders, as opposed to market orders, could be due to the effects of stale limit orders.

Table VIII

Portfolios Based on Limit Order Imbalances

At the end of each trading day, we sort all stocks with at least one limit order transaction within the brokerage according to their limit order imbalance (number of shares bought through limit orders minus number of shares sold through limit orders divided by total number of shares traded across all German stock exchanges) and assign them to one of three equally sized portfolios: the sell portfolio (PF1, the portfolio with the lowest, and typically negative, order imbalance), the hold portfolio (PF2), and the buy portfolio (PF3).

Part (i) reports the average order imbalance of the stocks in each of the three portfolios (PF1, PF2, PF3) and of a zero-cost portfolio (PF3-PF1), which is long in buy-portfolio stocks and short in sell-portfolio stocks. Part (ii) reports the corresponding average return (in percent), in excess of the equally weighted return across all German stocks for which Datastream provides total returns, measured from the close of the previous trading day to the close of the current trading day. All averages are reported for the day of portfolio formation (0) and the 5 days previous (−5 −1) and subsequent (1 5) to portfolio formation.

(i) Average buy pressure (in%)

	−5	−4	−3	−2	−1	0	1	2	3	4	5
PF1	0.036	0.040	0.035	0.018	−0.010	−1.439	−0.018	−0.003	0.007	0.020	0.029
	(0.012)	(0.010)	(0.011)	(0.011)	(0.011)	(0.040)	(0.010)	(0.011)	(0.010)	(0.012)	(0.012)
PF2	0.014	0.017	0.024	0.023	0.022	0.028	0.026	0.026	0.029	0.026	0.025
	(0.005)	(0.005)	(0.006)	(0.006)	(0.005)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)
PF3	0.133	0.137	0.126	0.144	0.199	1.826	0.182	0.148	0.111	0.127	0.112
	(0.014)	(0.014)	(0.013)	(0.015)	(0.015)	(0.051)	(0.017)	(0.015)	(0.015)	(0.014)	(0.014)
PF3–PF1	0.097	0.098	0.092	0.126	0.209	3.265	0.200	0.150	0.104	0.107	0.083
	(0.015)	(0.015)	(0.016)	(0.016)	(0.017)	(0.078)	(0.018)	(0.016)	(0.018)	(0.016)	(0.015)

(ii) Average returns (in%)

	−5	−4	−3	−2	−1	0	1	2	3	4	5
PF1	0.152	0.137	0.202	0.300	0.375	0.874	0.021	−0.096	−0.092	−0.029	−0.067
	(0.039)	(0.041)	(0.036)	(0.075)	(0.047)	(0.059)	(0.041)	(0.035)	(0.038)	(0.034)	(0.033)
PF2	0.013	0.054	0.066	0.143	0.182	0.153	0.009	−0.059	−0.079	−0.090	−0.053
	(0.037)	(0.036)	(0.040)	(0.041)	(0.042)	(0.050)	(0.037)	(0.036)	(0.035)	(0.033)	(0.035)
PF3	−0.016	−0.018	−0.012	−0.067	−0.094	−0.574	0.016	0.041	−0.014	−0.023	0.004
	(0.037)	(0.039)	(0.035)	(0.034)	(0.037)	(0.041)	(0.036)	(0.041)	(0.038)	(0.036)	(0.038)
PF3–PF1	−0.168	−0.156	−0.214	−0.367	−0.469	−1.448	−0.005	0.137	0.078	0.007	0.070
	(0.042)	(0.042)	(0.042)	(0.078)	(0.052)	(0.065)	(0.044)	(0.045)	(0.040)	(0.039)	(0.039)

IV. Conclusion

Retail investors act more similarly than would be expected by chance, even taking into account the fact that retail investors tend to move together because of short-sale constraints and unmonitored limit orders. The correlation among the sample investors is not driven by explicit brokerage advice, IPOs underwritten by the sample broker, or automatic investment plans offered by the sample broker. Therefore, the observed trades should be representative of the broader population of self-directed German retail investors. The positive correlation between observed market order imbalances and returns also points to this.

Retail trades appear to be mainly coordinated by speculative motives; retail investors move together into stock A primarily because they share the belief that stock A's price will appreciate more than that of other stocks, and not because stock A is part of a well-diversified savings portfolio. In this paper, we do not examine in detail what makes stock A a particularly attractive bet for retail investors—a difficult task that warrants future research. Rather, we document that, based on *market* order imbalances, retail speculators as a group behave as positive feedback traders, that is, they buy recent winners and sell recent losers. We also document that *limit* orders are executed as if they were placed by contrarians, which is likely because many of their orders become stale and are mechanically executed. Another possibility is that placed limit orders are different in some ways from market orders, or that those limit orders selected for execution by market orders are different. Future research with data on submitted, rather than just executed, limit orders could shed light on this interesting issue.

Finally, correlated retail trading appears to affect price formation through two channels. First, order imbalances based on market orders are positively correlated with contemporaneous as well as future returns. Although we cannot definitely rule out informed trading as an explanation for the return predictability, our results point to an explanation based on serially correlated price pressure. In other words, retail purchases today are associated with higher returns tomorrow because today's purchases are followed by more purchases, and thus price pressure, tomorrow. Second, limit order imbalances are negatively correlated with contemporaneous returns and positively correlated with future returns. Thus, limit order traders seem to be compensated for providing liquidity to other market participants.

Appendix

A. Simulation of Short-Sale Constraints

First, all stocks are sorted each period into deciles by the number of clients who hold the stock at the beginning of the period; holder decile 1 contains the most widely held stocks and holder decile 10 contains the least widely held stocks. To balance the number of transactions in each decile, exponentially more stocks are assigned to the deciles that contain less widely held stocks. Second, we simulate 1,000 data sets assuming that there is no correlated trading but that there are short-sale constraints. For each stock j in holder decile $\mathbb{S}_{j,t}$ and period t, we randomly draw the number of initial buyers, active holders, and sellers, assuming that they are binomially distributed with

$$\#InitialBuyers(j,t) \sim B\big(Total\#Investors(t) - \#Holders(j,t), \hat{p}_{newbuy}(j,t)\big)$$

$$\#ActiveHolders(j,t) \sim B\left(\#Holders(j,t), \hat{p}_{active}(j,t)\right)$$

$$\#Sellers(j,t) \sim B\left(\#ActiveHolders(j,t), \hat{p}_{sell}(j,t)\right),$$

where

$$\hat{p}_{newbuy}(j,t) = \frac{1}{\|\mathbb{S}_{j,t}\|} \sum_{k \in \mathbb{S}_{j,t}} \frac{\#InitialBuyers(k,t)}{Total\#Investors(t) - \#Holders(k,t)}$$

$$\hat{p}_{active}(j,t) = \frac{1}{\|\mathbb{S}_{j,t}\|} \sum_{k \in \mathbb{S}_{j,t}} \frac{\#RepeatBuyers(k,t) + \#Sellers(k,t)}{\#Holders(k,t)}$$

$$\hat{p}_{sell}(j,t) = \frac{1}{\|\mathbb{S}_{j,t}\|} \sum_{k \in \mathbb{S}_{j,t}} \frac{\#Sellers(k,t)}{\#RepeatBuyers(k,t) + \#Sellers(k,t)}$$

and # *Holders*(j, t) is the number of sample investors who hold stock j at the beginning of period t. Then, we calculate the average LSV measure across all stock-periods with at least two traders. The three probabilities vary considerably across holder deciles. For example, the ratio of actual initial buyers to potential initial buyers across all sample days is 0.013% for the most widely held stocks and 0.002% for the least widely held stocks; investors are more likely to buy into stocks that are already popular among their peers. Those who own less widely held stocks are more likely to trade them on any given day; the ratio of active holders to all holders is 3.5% for the least widely held stocks as opposed to 0.3% for the most widely held stocks. Moreover, owners of widely held stocks are more likely to increase their position than other owners; the ratio of owner-buyers to active owners in these stocks is 26.3% as opposed to 14.2% for the least widely held stocks.

REFERENCES

Amihud, Yacov, and Haim Mendelson, 1980, Dealership market: Market making with inventory, *Journal of Financial Economics* 8, 31–53.

Andrade, Sandro C., Charles Chang, and Mark S. Seasholes, 2007, Trading imbalances, predictable reversals, and cross-stock price pressure, *Journal of Financial Economics* (forthcoming).

Barber, Brad, and Terrance Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773–806.

Barber, Brad, and Terrance Odean, 2002, Online investors: Do the slow die first? *Review of Financial Studies* 15, 455–487.

Barber, Brad, Terrance Odean, and Ning Zhu, 2006, Do noise traders move markets? Working paper, University of California, Berkeley.

Barber, Brad M., Terrance Odean, and Ning Zhu, 2003, Systematic noise, Working paper, University of California, Berkeley.

Barberis, Nicholas C., and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161–199.

Campbell, John Y., Sanford J. Grossman, and Jiang Wang, 1993, Trading volume and serial correlation in stock returns, *Quarterly Journal of Economics* 108, 905–939.

Chordia, T., and A. Subrahmanyam, 2004, Order imbalance and individual stock returns, *Journal of Financial Economics* 72, 485–518.

Coval, Joshua D., and Tobias J. Moskowitz, 1999, Home bias at home: Local equity preference in domestic portfolios, *Journal of Finance* 54, 145–166.

Cutler, David M., James M. Poterba, and Lawrence H. Summers, 1989, What moves stock prices?, *Journal of Portfolio Management* 15, 4–12.

De Long, J. Bradford, Andrei Shleifer, Lawrence Summers, and Robert J. Waldmann, 1990a, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

De Long, J. Bradford, Andrei Shleifer, Lawrence Summers, and Robert J. Waldmann, 1990b, Positive feedback investment strategies and destabilizing rational speculation, *Journal of Finance* 45, 375–395.

Deutsches Aktieninstitut, 2003, Factbook 2002, Frankfurt am Main.

Dorn, Daniel, and Gur Huberman, 2005, Talk and action: What individual investors say and what they do, *Review of Finance* 9, 1–45.

Feng, Lei, and Mark S. Seasholes, 2004, Correlated trading and location, *Journal of Finance* 59, 2117–2144.

Frazzini, Andrea, and Owen Lamont, 2006, The earnings announcement premium and trading volume, Working paper, University of Chicago and Yale University.

Gilchrist, Simon, and Charles Himmelberg, 1998, Investment, fundamentals, and finance, Working paper 6652, National Bureau of Economic Research.

Glosten, Lawrence R., and Paul R. Milgrom, 1985, Bid, ask and transaction prices in a specialist market with heterogeneously informed traders, *Journal of Financial Economics* 14, 71–100.

Goetzmann, William N., and Massimo Massa, 2002, Daily momentum and contrarian behavior of index fund investors, *Journal of Financial and Quantitative Analysis* 37, 375–389.

Griffin, John M., Jeffrey Harris, and Selim Topaloglu, 2003, The dynamics of institutional and individual trading, *Journal of Finance* 58, 2285–2320.

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* 55, 43–67.

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.

Grossman, Sanford J., and Merton H. Miller, 1988, Liquidity and market structure, *Journal of Finance* 43, 617–637.

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

Hasbrouck, Joel, 1991, Measuring the information content of stock trades, *Journal of Finance* 46, 179–207.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

Huberman, Gur, and Werner Stanzl, 2004, Price manipulation and quasi-arbitrage, *Econometrica* 74, 1247–1276.

Hvidkjaer, Soeren, 2006, Small trades and the cross-section of stock returns, Working paper, University of Maryland.

Jackson, Andrew, 2003, The aggregate behaviour of individual investors, Working paper, London Business School.

Kandel, Eugene, and Neil D. Pearson, 1995, Differential interpretation of public signals and trade in speculative markets, *Journal of Political Economy* 103, 831–872.

Kaniel, Ron, Gideon Saar, and Sheridan Titman, 2008, Individual investor trading and stock returns, *Journal of Finance* 273–310.

Kumar, Alok, and Charles Lee, 2006, Retail investor sentiment and return comovements, *Journal of Finance* 61, 2451–2486.

Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1335.

Kyrolainen, Petri J., and Jukka Perttunen, 2002, Investors' activity and trading behavior, Working paper, University of Oulu.

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

Linnainmaa, Juhani, 2003, Who makes the limit order book? Implications for contrarian strategies, attention-grabbing hypothesis, and the disposition effect, Working paper, University of California at Los Angeles.

Love, Inessa, 2001, Estimating panel-data autoregressions, Working paper, Columbia University.

Malmendier, Ulrike, and Devin Shanthikumar, 2007, Are small investors naıve about incentives? *Journal of Financial Economics* 85, 457–489.

Mitchell, Mark L., and J. Harold Mulherin, 1994, The impact of public information on the stock market, *Journal of Finance* 49, 923–950.

New York Stock Exchange, 2001, U.S. shareholders and online trading, New York.

Nofsinger, John R., and Richard W. Sias, 1999, Herding and feedback trading by institutional and individual investors, *Journal of Finance* 54, 2263–2295.

Shleifer, Andrei, and Robert Vishny, 1997, The limits to arbitrage, *Journal of Finance* 52, 35–55.

Stoll, Hans, 1978, The supply of dealer services in securities markets, *Journal of Finance* 33, 1133–1151.

Van Steenis, Huw, and Christian Ossig, 2000, European online investor, JP Morgan Equity Research Report.

Wermers, Russ, 1999, Mutual fund herding and the impact on stock prices, *Journal of Finance* 54, 581–622.

Williams, Rick, 2000, A note on robust variance estimation for cluster-correlated data, *Biometrics* 56, 645–646.

Womack, Kent L., 1996, Do brokerage analysts' recommendations have investment value? *Journal of Finance* 51, 137–168.

Wylie, Sam, 2005, Fund manager herding: A test of the accuracy of empirical results using U.K. data, *Journal of Business* 78, 381–403.