Trading Styles and Trading Volume

Stefan Nagel*

Stanford University and NBER

August 2005

I thank Viral Acharya, Malcolm Baker, Nick Barberis, Markus Brunnermeier, Elroy Dimson, James Dow, Wayne Ferson, Francisco Gomes, Ilan Guedj, Denis Gromb, Jon Lewellen, Chris Malloy, Stijn van Nieuwerburgh, Jannette Papastaikoudi, Andrei Shleifer, Jeremy Stein, Hongjun Yan, and seminar participants at Boston College, Columbia, Dartmouth, Harvard Business School, London Business School, London School of Economics, University of Mannheim, MIT, NYU, Stanford, and Yale for comments. Thanks also to Chris Malloy and I/B/E/S for providing data on analyst forecast dispersion.

*Stanford University, Graduate School of Business, 518 Memorial Way, Stanford, CA 94304, USA. Nagel_Stefan@gsb.stanford.edu; http://faculty-gsb.stanford.edu/nagel; Phone 650-724-9762, Fax 650-725-7979.

Trading Styles and Trading Volume

ABSTRACT. Many investors in equity markets follow trading styles based on stock characteristics like firm size or momentum. Because stocks' characteristics change over time, style investors must rebalance their portfolios to avoid style drift. In this paper, I analyze how the rebalancing trades of style traders affect the cross-section of trading volume. Specifically, I estimate, using mutual fund stock holdings data, how the propensity to sell a stock depends on recent changes in stock characteristics. It turns out to be most strongly related to changes in firm size, value, and momentum characteristics. The fitted values of the propensity-to-sell model explain a significant portion of the cross-sectional variation in total trading volume, controlling for other determinants of volume and using instrumental variables to deal with the endogeneity of stock characteristics. The results suggest that style investing is an important source of trading volume.

Keywords: Trading Volume; Style Investing; Mutual Funds

JEL classification: G10, G20

1. Introduction

One striking fact about trading in equity markets is the widespread specialization of investors into various kinds of trading styles. Some professional traders focus on technical analysis and quantitative trading strategies; mutual funds specialize into value or growth, small or large-cap stocks; and retail investors, too, often follow styles such as momentum or contrarian trading strategies, for example (see, e.g., Brown and Goetzmann 1997; Fung and Hsieh 1997; Grinblatt and Keloharju 2001; Chan et al. 2002). The common thread of these "styles" is that investors aim to hold stocks that have a few common pre-defined characteristics, such as, say, positive momentum or a high book-to-market ratio.

Style trading has interesting implications for trading volume. When stocks' characteristics change over time and some stocks drift away from the style target, style investors have to trade to discard those stocks that are out-of-line. The more volatile the characteristic the style is based on and the narrower investors' style focus, the higher the trading volume that is required to maintain a style-consistent portfolio. Given the apparent pervasiveness of style investing in equity markets, these rebalancing trades of style investors could potentially account for a substantial portion of trading volume. If so, we can hope for two kinds of insights from an investigation of trading styles: From a purely empirical perspective, it could improve our understanding of the determinants of trading volume. From a theoretical perspective, it could point to trading styles as an important channel for trading activity, which carries implications for why investors trade (so much)—a question that is often seen as a puzzle by researchers, including Black (1986), Ross (1989), and Dow and Gorton (1997).

To shed light on this issue, I analyze empirically the implications of style trading for the cross-section of trading volume. The main challenge is to find ways to quantify the extent of style trading and its consequences for trading volume. In this respect, the key innovation in this paper is the estimation, using mutual fund portfolio holdings data, of how investors' propensity to sell a stock depends on changes in certain stock characteristics. The idea is that if many funds follow trading styles based on these characteristics, then the more a stock's characteristics change, the more likely it is that the previous owners will sell because the stock no longer fits into their style category. To take a simple example, when a stock's price-to-earnings ratio is low, it is likely to be owned by investors following a "value" style. If the price-to-earnings ratio increases, value investors are likely to sell to "growth" investors. I estimate the relationship between the propensity to sell and changes in stock characteristics with a multivariate linear probability model. The estimated model shows that a one cross-sectional standard deviation change in the momentum (12-month past returns), value (sales-to-price) and size (market capitalization) characteristic each lead to a marginal increase in the propensity to sell by about 3%, measured over an annual period. One-standard deviation changes in long-term growth rate forecasts and past 36-months returns each contribute a further 2%. These characteristics appear to be the most important drivers of style-related trading volume.

The linear probability model provides a convenient way to aggregate the effects across different characteristics. For each stock in each year, the observed changes in all of the characteristics can be used to calculate the fitted propensity to sell. This fitted propensity to sell should be correlated with differences in trading volume across stocks. Indeed, to the extent that not only mutual funds, but also other investors follow trading styles, the propensity-to-sell measure should be correlated with cross-sectional differences not only in the volume of trading by mutual funds, but also in overall trading volume. For NYSE and AMEX stocks I find that a one cross-sectional standard deviation difference in the propensity to sell measure is associated with a difference in annual turnover of around 20% to 30%. For Nasdaq stocks it is about twice that number, consistent with the fact that the recorded trading volume on Nasdaq is about twice as high as on NYSE and AMEX due to the different market structure. Compared with the average turnover of about 87% and 174% per year for NYSE/AMEX and Nasdaq stocks, respectively, this is a sizeable magnitude. Apparently, a considerable part of trading volume in equity markets is indeed driven by the rebalancing needs of style traders in response to changes in stock characteristics. In fact, this estimate might be conservative, because the methodology in this paper is not designed to capture the effects of trading rules operating at high frequency—such as various technical trading rules, for example.

The results are robust to controlling for other known determinants of cross-sectional differences in trading volume, such as volatility, market capitalization, and stock price. One concern that one might have about these results is that trading causes prices to change which in turn leads to changes in price-based stock characteristics such as book-to-market or momentum. Under this alternative view, it would be volume that drives changes in characteristics rather than the other way round. Yet, the results still hold up when the propensity-to-sell measure based on all stock characteristics is instrumented with a propensity-to-sell measure calculated using only stock characteristics that are not price-based, such as book leverage or book return on assets, for example. Hence, the correlation between trading volume and the propensity-to-sell measure is not driven by reverse causality from volume to prices and price-dependent characteristics.

In related empirical work researchers have shown that there is abnormal trading volume following certain technical trading signals, which may indicate the presence of traders with technical trading rules. Chu and Osler (2004) show that trading volume is abnormally high following "head-and-shoulders" patterns in stock prices. Huddart, Lang, and Yetman (2003) find that there are pikes in trading volume when a stock's price leaves a prior "trading range". My analysis in this paper is also related to work by Lo and Wang (2000). Lo and Wang show that there is a factor structure in trading volume when investors switch their investments between separating funds, i.e. baskets of stocks. In their setting stocks do not switch between baskets. This paper, too, is based on the notion that investors hold baskets of stocks, defined by some stock characteristics, but it focuses on the trading volume effects when stocks, not investors, switch between baskets. Of course, style-switching by investors, for example as in Barberis and Shleifer (2003), could lead to additional trading volume. However, while the rebalancing trades of investors with fixed styles can be related empirically to changes in stock characteristics, it is not clear how one could isolate empirically the trading volume effects of style-switching. Hence my focus in this paper on the trades arising from fixed trading styles.

Existing theoretical work is largely silent on trading styles and their implications for volume. In one strand of the theoretical volume literature, trading arises from differential or asymmetric information, coupled with trading for liquidity or hedging reasons (e.g., Admati and Pfleiderer 1988; Wang 1994; He and Wang 1995). Because trading styles are based mostly on public information, style trading does not readily fit into this framework. In a different class of models, trading arises from differences in beliefs (Varian 1989; Harris and Raviv 1993; Kandel and Pearson 1995). Differences in trading styles could reflect differences in beliefs to the extent that heterogeneity in trading styles arises because investors use different data sets, differ in modeling techniques and sophistication, employ different heuristics to simplify the decision problem, and therefore come up with different trading rules. In Mullainathan (2002) and Hong and Stein (2003), for example, investors use such simplified models to form their beliefs. This view also fits with the notion in Black (1986) that some investors trade one "noise" because they use simple rules of thumb. However, the pervasiveness of trading styles among institutional investors may also have to do with industrial organization and agency issues (Massa 2000; Mamaysky and Spiegel 2002; Barberis and Shleifer 2003), which are issues that the trading volume literature typically abstracts from (with the exception of Dow and Gorton 1997). The evidence I present in this paper indicates that these institutional arrangements in money management may hold clues about the origins of trading volume.

2. Estimating the trading volume effects of style investing

2.1 The propensity-to-sell measure

The degree to which style investing drives trading volume depends on how many style traders there are, how strongly their portfolios are focused on a certain style, and how volatile a stock's style-related characteristics are. Existing work attempting to measure the extent to which investors engage in style investing examines how strongly investors cluster in terms of returns (Brown and Goetzmann 1997) or in terms of the characteristics of stocks in their portfolios (Wermers 2002). However, none of these style measures allows a quantitative assessment of how changes in stock characteristics translate into rebalancing needs, and therefore trading volume, for style traders.

My empirical strategy uses the fact that, on average, a stock that experiences a change in a style-relevant characteristic tends to move away from its current owner's style target (e.g., if a "value" stock is owned by a "value" investor, mean-reversion in the characteristic implies that it will tend to drift towards "growth"). Hence, the bigger the change in the characteristic, the more likely it is that the current owner will sell it. Of course, this is unlikely to be true in every single case, but it should hold on average. The more style traders there are, and the stronger there style focus, the greater should be the average effect of a change in characteristics on the propensity to sell a stock.

Following this intuition, I use mutual fund portfolio holdings data to estimate how the typical mutual fund's propensity to sell a stock depends on changes in this stock's characteristics. More precisely, to model how mutual funds j's selling decisions during period t are related to changes in stock i's characteristics during the same period, I use a linear probability model

$$P(s_{jit} = 1 \mid |\Delta c_{it}|) = \alpha_t + \beta' |\Delta c_{it}| \,, \qquad (1)$$

where the indicator variable s_{ijt} is a Bernoulli random variable that equals one if all of mutual fund j's holdings of stock i are sold during period t and zero if not. The vector Δc_{it} represents the changes in stock i's characteristics during period t and β is a column vector of regression coefficients. I estimate this model with OLS. Fitted values, $\hat{s}_{it}$, then provide the probability, conditional on Δc_{it}, that a current owner of stock i will sell it. A disadvantage of linear probability models can be that some fitted values are higher than one or smaller than zero. However, it turns out that in this application here this never happens. Because coefficients are easier to interpret in a linear model than in a Probit or Logit model, I use a linear model.

In principle, one could also estimate a relationship similar to Equation (1) with each stock's total trading volume as dependent variable. I use a two-step approach—first estimating (1) with mutual fund portfolio holdings data and then using the fitted value to explain total trading volume—

because of several reasons. Most importantly, by focusing on quarterly portfolio holdings of mutual funds in the first step, we are looking at trades for which we can largely rule out alternative trading motivations such as dynamic hedging of derivative exposures, for example. Moreover, from the prior literature, we know that mutual funds tend to follow styles. These conditions make it more plausible that what the regressions pick up is indeed related to style investing, and not some other reason for trade. Testing whether the fitted values from the first step have explanatory power for total trading volume in the second step then provides a useful out-of-sample check (given that the share of outstanding equity held by mutual funds is relatively small), showing whether the selling propensities estimated for mutual funds have any predictive power for what other investors tend to do. Also, having aggregated the changes in all the different characteristics into one fitted value, it is then easier to introduce control variables in the second step (without running into multicollinearity problems) and to use instrumental variables.

Instead of using the propensity to sell to estimate equation (1), I could also use the propensity to buy. However, since funds have a much larger universe to choose from in their buying decisions compared with their selling decisions (the portfolio holdings data do not include short positions, but they are rare for mutual funds anyway, as shown by Almazan, Brown, Carlson, and Chapman 2004). The propensity to buy is therefore much smaller on average and harder to estimate than the propensity to sell.

2.2 Data

Estimation of the propensity to sell requires portfolio holdings data, which I obtain from the Thomson Financial Mutual Funds (Spectrum) Database. The database contains quarterly stock-by-stock positions of most U.S. mutual funds, and is explained is explained in great detail in Wermers (1999). The timing of holdings is based on the report date (RDATE) in the Thomson database. For those funds that do not report each quarter, I fill in the missing quarter by bringing forward the latest available (stale) holdings report, but using current prices. The vast majority of funds report data at least at the semiannual frequency (see, also, Wermers 1999). When RDATE and vintage date (FDATE) do not agree, Thomson applies a split-adjustment that I undo using split-adjustment factors from CRSP to make the data comparable to other observations with the same RDATE.

I match these quarterly holdings observations with data on stock characteristics from CRSP, COMPUSTAT and I/B/E/S. In principle, one would want to select the stock characteristics that mutual funds are known to use as a basis for their trading styles. There is some guidance from the existing literature as to what these characteristics are. Brown and Goetzmann (1997), Chan et al. (2002), and Wermers (2002) show that mutual funds tend to cluster on firm size, measures related to growth or glamour (e.g., book-to-market) and momentum; Hotchkiss and Lawrence (2003) find that institutional investors differ in their preferences for dividend yield. However, an exhaustive search has not yet been conducted in the literature. For this reason, I consider a wide range of characteristics and let the data speak as to their relevance for mutual funds' trading styles.

My initial list of candidate characteristics is drawn up based on the following principles. First, they should be plausibly related to firm value or risk. The idea here is that investors might most likely condition their trading rules on such characteristics. Second, the characteristics should be continuous variables. This rules out discrete variables such as analyst recommendations or index membership. Third, the characteristics should have sufficient time-variation. Only relatively volatile characteristics are likely to be sources of rule-driven trading volume. For example, while it might seem plausible that some investors select firms based on R&D expenses, most firms have quite persistent ratios of R&D expenses to sales, making it an unlikely candidate to explain trading volume. The same is true for many other accounting variables.

Based on this reasoning, I select the following characteristics: Momentum (Return(-12:-1) = past 12-month returns; Return(-3:-1) = past 3-month returns), value (or "glamour") characteristics

(Return(-36:-1) = past 36-month return; Sales/MarketCap[1]), profitability (Earnings/Assets), leverage (Liabilities/MarketCap), firm liquidity (Cash/CurrLiab = cash and short-term investments/current liabilities), expected growth (LTGrowthForecast = I/B/E/S median long term growth rate forecast), yield (DividendYield), risk measures (Beta = past 36-month beta on the value-weighted CRSP index; Volatility = one-year sum of squared daily returns), and firm size (MarketCap). In the end, any prior selection of characteristics involves some arbitrary element, but the results are unlikely to be sensitive to the exact definition of the characteristics set. It seems hard to think of other firm characteristics that would not be correlated with the chosen ones. Two additional characteristics (book liabilities/book assets and the one-year growth rate in sales) will be used as instruments in some of the tests below. For an observation to be included in the tests in a given quarter, availability of valid observations on all characteristics, including the two additional instruments, is required. Because investors are likely to select stocks based on the relative rank rather than the absolute level of characteristics, and because the cross-sectional distributions of some of the raw characteristics are not well-behaved, I transform all characteristics into percentile ranks each quarter.

Because one-year changes are required for all characteristics, and I/B/E/S long-term growth forecasts are not available for a sufficiently large number of stocks before 1983, the sample period starts in the first quarter of 1984. It ends in the fourth quarter of 2002. The sample includes NYSE, AMEX and Nasdaq stocks, excluding primes, scores, and ADRs. Tests using trading volume reported on CRSP, however, are run separately for NYSE/AMEX and for Nasdaq stocks, because trading volume figures are not directly comparable due to different market structures (see Atkins and Dyl 1997).

In each quarter and for every stock, I use the transformed characteristics to measure changes in characteristics (Δc_{it})—the inputs for the linear probability model—over overlapping annual intervals. There are two reasons for this choice of interval: First, some characteristics based on accounting data

---

[1] I use sales/price and not the perhaps more popular book-to-market to avoid having to eliminate observations with negative book values. Towards the end of the 1990s there is a fairly large number of firms with negative book values, for which the book-to-market ratio is not a meaningful valuation ratio.

change only once per year. Shorter observation intervals would not make much sense for these characteristics. Second, for some funds, the portfolio holdings data are reported at semi-annual or even annual frequency only. Annual observation intervals thus circumvent complications arising from stale holdings information. Correspondingly, the selling indicator variable, s_{ijt}, is also measured over annual intervals. More precisely, it equals one if fund j holds zero shares of stock j at time t, but held more than zero shares four quarters earlier. It equals zero, if fund j holds more than zero shares at both time t and four quarters earlier. It is set to missing for fund-stock observations where fund j did not hold a position in stock i four quarters earlier.

2.3 Summary statistics

Table 1 presents summary statistics for the sample of stocks and mutual funds used in this study. These statistics are time-averaged cross-sectional estimates. As shown in the bottom line, the data requirements leave, on average, 1,730 stocks in the sample, which is about 30**%** of all eligible stocks on CRSP. Yet, since these stocks tend to be large stocks, they account for about 75% of market capitalization. The average number of mutual funds in the Thomson Financial database that hold at least one of these stocks with valid data is 3,788, and the median fund holds 39 stocks satisfying the data requirements, while the mean is 25.

The first set of rows in Table 1 shows statistics on the cross-sectional distribution of mutual funds' mean portfolio characteristics. Mean portfolio characteristics are the value-weighted average, using portfolio weights from the quarterly holdings snapshots, of the characteristics of the stocks held by a fund. The cross-sectional distribution of these mean portfolio characteristics reveals whether mutual funds in aggregate have preferences for certain characteristics. To judge the magnitudes, recall that all of the underlying characteristics are transformed to percentile ranks (ranging from 0 to 99). As the first block of rows shows, the mean and median fund seem to have some preference for stocks with high past returns, growth stocks (low Sales/MarketCap), high dividend yield, low Volatility, and large size. The 5^{th} and 95^{th} percentile values show that the most dispersion in funds' characteristics exposure is with respect to forecasted long-term growth and dividend yield. Such dispersion hints at the presence of heterogeneous trading styles. Alternatively, the dispersion could simply reflect the fact that funds are randomly switching between different groups of stocks. Whether funds persistently stick to trading styles cannot be determined from dispersion alone.

The matrix in the second block shows cross-sectional correlation estimates for stock characteristics, c_{it}, (upper triangular part) and funds' mean portfolio characteristics. The most salient patterns are that return-based characteristics have, not surprisingly, strongly positive correlation; stocks with high Sales/MarketCap tend to have low profitability (Earnings/Assets), high leverage (Liabilities/MarketCap), and low expected growth (LTGrowthForecast); firms with high profitability have high leverage; and high dividend yield stocks tend to have low forecasted growth and low volatility and beta. The multivariate version of the linear probability model that I estimate below accounts appropriately for these correlations. A pairwise comparison of the correlations of funds' mean characteristics and of the correlations of characteristics themselves shows that they are usually quite similar. To provide a specific example, the typical high yield fund also holds stocks with low volatility (corr. –0.85), just as the typical high yield stock tends to have low volatility (corr. –0.65).

Table 2 provides some simple summary statistics for annual trading volume on NYSE/AMEX and Nasdaq, as well as a proxy for mutual fund trading volume constructed from the quarterly mutual fund holdings data. This is the data that the propensity-to-sell measure will be asked to explain. The table shows trading volume measured as turnover. Turnover is calculated as the number of shares traded during a month divided by the average of the number of shares outstanding at the beginning and at the end of the month. This monthly turnover is then averaged summed up over 12-month periods. Mutual fund selling volume is calculated from the portfolio holdings data. For each stock i, I compare mutual funds' holdings at time t to those 12 months earlier. For each fund with non-zero holdings 12 months earlier, sales are defined as the reduction in the number of shares (split-adjusted) between these two dates (zero if no reduction). Summing all sales across funds, selling volume is then defined as the number of shares of stock i sold by all mutual funds, divided by the number of shares held 12 months earlier.

As can be seen in the table, the (equal-weighted) mean turnover on Nasdaq (174%) is about twice as large as on NYSE and AMEX (87%). This discrepancy reflects the fact that Nasdaq is a dealer market where trading volume can be double-counted. Some researchers therefore divide Nasdaq turnover by two to make it comparable to NYSE and AMEX trading volume.[2] In this paper, instead, I run separate regressions for Nasdaq and NYSE/AMEX data. With 42%, the turnover of mutual funds—in terms of selling volume—is considerably lower. It is also lower than the turnover of mutual fund portfolios reported, for example, in the CRSP mutual funds database, which is around 80% - 90% per year over the same sample period. This partly reflects the fact that the mutual fund trading volume proxy is constructed by comparing holdings at annual intervals, which misses intra-year fluctuations in holdings and therefore a considerable part of higher-frequency trading volume.

3. Results

3.1 Propensity-to-sell estimates

Table 3 presents three types of information. The first column reports the mean absolute change in each stock characteristic. This is a measure for the volatility of each characteristic. The second column shows the coefficient estimates of the linear probability model, Equation (1), when it is run as a univariate model, with only one characteristic as an explanatory variable. The third column presents the coefficient estimates of the multivariate version of the model, where all characteristics are included jointly (intercepts are omitted in the table). To make the coefficient estimates easily comparable, changes in each stock characteristic are standardized by their average cross-sectional standard deviation before running the regressions. Both the characteristics volatility in the first column

[2] See Atkins and Dyl (1997) for a detailed discussion of this practice.

and the slope coefficient in the second and third columns are first estimated for each cross-section. The table shows the time-averaged estimates. For the slope coefficient estimates, one could draw inference by calculating the standard error from their variation across time. However, it turns out that these standard errors are tiny and negligible for all slope coefficient estimates. Therefore, to reduce clutter, they are not reported in the table.

Focusing on the first column, it can be seen that the short-term momentum characteristics Return(-12:-1) and Return(-3:-1) are the most volatile characteristics. Their mean absolute change is close to 40 (For comparison, recall that characteristics are expressed as percentile ranks ranging from 0 to 99). Not surprisingly, the least volatile characteristic is size (MarketCap) with a mean absolute change of 8.26. With mean change of 9.42, the dividend yield is also quite stable over time.

The univariate regression slopes from the linear probability model in column 2 show that one-standard deviation changes in MarketCap, Liabilities/MarketCap, and Sales/MarketCap are associated with the highest increases in mutual funds' propensity-to-sell. For example, a one-standard deviation change in Sales/MarketCap leads to an increase in the propensity to sell by 7.6%. The characteristic with the smallest impact is Cash/CurrentLiab with a slope coefficient of 1.8%.

To check whether a linear model provides a sufficiently good approximation, Figure 1 provides some non-parametric estimates of the relationship between changes in characteristics and the propensity-to-sell. I use a simple binning estimator. Stocks are binned by the standardized change in their characteristic with a bin width of 0.01 standard deviations. Within each bin, I then calculate the equal-weighted mean of s_{ijt}. The time-averages of these quarterly means are shown as solid lines, plotted against the (standardized) change in characteristics. The dotted lines represent two-standard-error bands, calculated from the time-variation of the quarterly means. It is apparent from these graphs that for the three return-based characteristics, there is not much increase in the propensity to sell for small changes in characteristics, only for larger changes the slope is steeper. In contrast, for size (MarketCap) the propensity to sell increases quickly with small changes size and then the curve flattens out a bit. But otherwise the graphs show that there is not too much non-linearity, which suggests that the linear model should be a reasonably good approximation.

The coefficient estimates from the univariate model, however, cannot be used to aggregate the propensity to sell across characteristics, because this would ignore correlation among characteristics. The multivariate model in column 3 takes this correlation into account. As a result, the slopes in the multivariate model are generally lower than in the univariate model. For example, the coefficient on MarketCap drops from 6.51% to 3.09% but it remains among the characteristics with the highest slopes. The coefficient on Liabilities/MarketCap drops from 7.30% to 0.73%, which indicates that most of the univariate slope is driven by correlation with changes in other characteristics, but holding those fixed, changes in Liabilities/MarketCap have little impact on the propensity to sell. In the multivariate model, the most important characteristics with the highest slopes are MarketCap, Sales/MarketCap, and Return(-12:-1). Thus, the size, value, and momentum characteristics that the previous literature on mutual fund styles has focused on also turn out to be most relevant here in explaining mutual funds' selling decisions.

While those three characteristics are the most important drivers of the propensity to sell, which is the main focus in this paper, size appears to be the single most important one in terms of how narrowly mutual funds focus their portfolios. A one standard deviation move in MarketCap is much smaller than for Sales/MarketCap or Return(-12:-1), which means that mutual funds tolerate much less dispersion with respect to size in their portfolios compared with value or momentum characteristics. It is also useful to keep in mind that the fact that MarketCap, Sales/MarketCap, and Return(-12:-1) turn out to be the most important drivers of the propensity to sell does not mean that funds actually use exactly those characteristics for their portfolio decisions. They might instead use other characteristics that are not included here, but that are highly correlated with MarketCap, Sales/MarketCap, and Return(-12:-1). For example, many technical trading signals are likely to be correlated with the Return(-12:-1) characteristic. A stock that switches from being a loser to being a winner stock is also likely to break through "trendlines" and experience changes in other technical indicators. For the purposes of this paper that's fine. In that case, the characteristics used here pick up the effects of these other unobserved characteristics.

3.2 Explaining the cross-section of trading volume

I now explore whether the propensity-to-sell estimates from the fitted model help explain the cross-section of total trading volume. The estimation of the propensity-to-sell model has focused on mutual funds, because this is the only group of investors for which comprehensive portfolio holdings data is available over a long sample period. However, other classes of investors are also likely to engage in style trading. To the extent that their style trading is similar to that of mutual funds, the propensity to sell estimates for mutual funds should be useful to predict overall trading volume in stock markets, not just mutual fund trading volume.

Table 4 presents the results of pooled OLS panel regressions of total turnover on each stock's fitted propensity to sell, which is calculated using the coefficient estimates from the multivariate model shown in Table 3. In addition to the contemporaneous propensity to sell, $\hat{s}_{it}$, I also use its first lag, $\hat{s}_{it-1}$, as an explanatory variable. The reason is that investors might react with some delay to changes in a stock's characteristics. Using only $\hat{s}_{it}$, which is based on contemporaneous changes in characteristics, would ignore the trading volume that arises with a lag. In addition to $\hat{s}_{it}$ and $\hat{s}_{it-1}$, some specifications also contain stock return volatility (the sum of squared daily returns over the contemporaneous annual interval), firm size (beginning of period log of market capitalization), and stock price (beginning of period log of price) as control variables. Lo and Wang (2000) show that these variables are the important determinants of trading volume in the cross-section. For ease of interpretation, all explanatory variables, including $\hat{s}_{it}$ and $\hat{s}_{it-1}$, are standardized to unit cross-sectional standard deviations and zero mean within each quarterly cross-section. In addition, all specifications also contain time-period dummies (coefficients not reported).

Because there is commonality in turnover, as shown by Lo and Wang (2000) and Cremers and Mei (2004), the residuals in these regressions are likely to have substantial cross-sectional correlation.

Moreover, since turnover is measured quarterly using overlapping annual intervals, there is substantial autocorrelation. To get accurate standard errors in the presence of such cross-sectional and time-series dependence in residuals, I use a moving-block bootstrap. It involves drawing entire blocks of adjacent quarterly cross-sections to construct bootstrap samples. The method is described in detail in the appendix.

Panel A presents the results for NYSE and AMEX stocks. Regression A.1 contains only $\hat{s}_{it}$ and $\hat{s}_{it-1}$ as explanatory variables. The point estimate for the coefficients on $\hat{s}_{it}$ and $\hat{s}_{it-1}$ of 0.18 (std.err. 0.02) and 0.08 (std.err. 0.02), respectively, imply that a one-standard deviation increase in $\hat{s}_{it}$ is associated with an increase in trading volume of 18% contemporaneously and an additional 8% at a one-year lag. The adjusted R^2 is about 10%. Compared with a mean turnover for NYSE and AMEX stocks of 87% per year and a cross-sectional standard deviation of 65% per year, these coefficient estimates, taken at face value, indicate a relatively large impact of style trading on turnover. In particular, there are other sources of trading activity besides style trading and so one would not expect style trading to account for all trading volume. Some of these other sources of trading activity may also be correlated with changes in stock characteristics and hence with $\hat{s}_{it}$. Specification A.2 is an attempt to control for such other sources of trading activity. It includes volatility, which can be seen as a proxy for information-based trading (Karpoff 1987). The inclusion of size is motivated by the fact that trading volume of larger firms is more likely to be affected by index arbitrage and program trading. Including the level of the stock price helps controlling for thin trading of stocks with extremely low prices. Interestingly, including these control variables does not have a major effect on the estimates for the coefficients on $\hat{s}_{it}$ and $\hat{s}_{it-1}$. With 0.14 (std.err. 0.01) and 0.10 (std.err. 0.01) they are still large. Specification A.3 includes a 12-month average of the Diether, Malloy, and Scherbina (2002) analyst forecast dispersion measure. The idea here is that analyst forecast dispersion might proxy for differences in opinion among investors other than those that manifest itself as different trading styles. The results show that it is indeed positively related to trading activity, but with an economically small coefficient estimate of

0.02 (std.err. 0.01). It leads to a further small drop in the coefficients on $\hat{s}_{it}$ and $\hat{s}_{it-1}$ to 0.12 and 0.09, respectively.

Panel B repeats these regressions for Nasdaq stocks. The results are remarkably similar to the NYSE/AMEX ones. Apart from the fact that coefficients are all a bit more than twice the NYSE/AMEX coefficients, there is little difference between Panel A and Panel B. This difference in magnitudes simply reflects the fact that trades on Nasdaq are double-counted, compared with NYSE/AMEX.

In Panel C, for mutual fund trading volume, the coefficients on $\hat{s}_{it}$ and $\hat{s}_{it-1}$ are smaller than in Panel A. For example, in specification C.1, the coefficient estimates of 0.06 and 0.02, respectively, are about one third of the values in Panel A, specification A.1. Most likely, this difference reflects the fact that the mutual fund trading volume proxy is constructed by comparing portfolio holdings snapshots at annual intervals. This misses intra-year trades if a fund purchases and sells shares of the same stock multiple times during a year. Apart from the fact that coefficient estimates are generally closer to zero, the results in for mutual fund trading volume are similar to those for total trading volume in Panels B and C.

One potential problem with these OLS regressions in Table 4 is that the propensity-to-sell measures on the right-hand side are endogenous. In particular, one might be concerned that what the regressions pick up is not that changes in stock characteristics make style investors trade, but that trade occurring for other reasons causes prices to change, which in turn leads to changes in stock characteristics that are based on price, such as past returns or the Sales/MarketCap ratio, for example. In other words, there could be a reverse causality problem. I address this concern by instrumenting $\hat{s}_{it}$ with the propensity-to-sell measure calculated in the same way, but using only characteristics that are not based on price. I also exclude the analyst growth long-term growth forecast characteristics, because analysts might use changes in prices as a signal to update their forecasts. Specifically, I redo the estimation of the multivariate linear probability model using only Earnings/Assets, Cash/CurrLiab, one-year sales growth (as a replacement the long-term analyst growth forecasts), Liabilities/Assets (which replaces market leverage). I then use the fitted value of this modified propensity to sell and its lag as instruments, denoted $\hat{s}_{it}$* and $\hat{s}_{it-1}$*. The instruments are correlated with $\hat{s}_{it}$ and $\hat{s}_{it-1}$, but they are not subject to the reverse causality critique and should therefore be valid instruments.

This instrumental variables strategy should address the concerns about reverse causality in the relationship between trading volume and $\hat{s}_{it}$. It is important to keep in mind, though, that this does not fully resolve all potential endogeneity issues in these regressions. For example, trading volume and volatility are simultaneously determined in the trading process. Since volatility is endogenous, the coefficient estimate on volatility therefore does not have a structural interpretation, and this could also have some effect on the coefficient estimates on $\hat{s}_{it}$ and $\hat{s}_{it-1}$. Ideally, one would like to have controls for the structural causes of trading volume unrelated to style trading, such as information flow, for example. However, lacking a structural model and suitable instruments for the latent variables causing volatility and volume there is little more that can be done.

Table 5 presents the results of the two-stage least squared (2SLS), using $\hat{s}_{it}$* and $\hat{s}_{it-1}$* as instruments for $\hat{s}_{it}$ and $\hat{s}_{it-1}$. The specification of the regression is similar to models A.2, B.2, and C.2 in Table 4 (Unreported results show that including forecast dispersion here, too, has little effect). To get the standard errors right in the presence of cross-sectional and auto-correlation of the residuals, I use a moving-block bootstrap in similar fashion as for the OLS regressions above. Panel A shows estimates for the sample with NYSE and AMEX stocks. As can be seen in the table, the point estimates of the coefficients on $\hat{s}_{it}$ and $\hat{s}_{it-1}$ (0.17 and 0.16, respectively) are somewhat larger than with OLS. However, standard errors are larger, too (0.04 and 0.03, respectively). As a result, the 2SLS point estimates are less than two standard errors away from the OLS point estimates. So there is a lot of uncertainty about whether they really are different. For Nasdaq stocks in Panel B coefficients are also larger than in the OLS regressions. For mutual fund volume in Panel C, the point estimates are almost identical to those with OLS in Table 4. Overall, the 2SLS regressions confirm that $\hat{s}_{it}$ and $\hat{s}_{it-1}$ have a strong influence on the cross-section of trading volume. The 2SLS results do not give reason to suspect that the positive relationship between $\hat{s}_{it}$ and $\hat{s}_{it-1}$ discovered in the OLS regressions is driven by reverse causality.

3.3 Subperiod results and out-of-sample tests

To check whether style trading had an impact on trading volume throughout the sample period, Table 6 presents subperiod results, where the sample is split into the 1985 – 1993 and 1994 – 2002 periods. In addition, Table 6 also presents out-of-sample results where the multivariate linear probability model for the propensity-to-sell measure is estimated in the first subperiod, and the thus estimated coefficients are used in the second subperiod to calculate the fitted values, $\hat{s}_{it}$, used in the panel regressions to explain the cross-section of trading volume. This out-of-sample test should address concerns that the linear probability model picks up some spurious, sample-specific correlations between characteristics changes and mutual fund trades. Because mutual fund trades are part of overall trading volume, this could then lead to a spurious between $\hat{s}_{it}$ and total trading volume.

As can be seen in Table 6, the estimates for the coefficients on $\hat{s}_{it}$ and $\hat{s}_{it-1}$ are somewhat larger in the second subperiod (1994 – 2002), both for NYSE and AMEX stocks (Panel A) and for Nasdaq stocks (Panel B). For mutual fund trading volume in Panel C, the coefficient estimates are almost identical across subperiods. A potential explanation for these results is that mutual funds hold a much larger fraction of the total market in the second subperiod. Based on the portfolio holdings data, their share of total market equity of the stocks in my sample has increased from about 5% in 1985 to about 23% at the end of 2002. If style trading is more prevalent among mutual funds than among other investors, or if the estimated propensity-to-sell model would be different for other investors, then, as mutual funds' share of the market and total trading volume goes up over time, one would expect the coefficient estimates on $\hat{s}_{it}$ and $\hat{s}_{it-1}$ to change in the regressions with total trading volume (Panels A and B), but not the in regressions with mutual fund trading volume (Panel C).

Interestingly, the coefficient estimates in the out-of-sample regressions (A.3, B.3, and C.3) are similar to those in the in-sample regressions for the same subperiod 1994 – 2002 (A.2, B.2, and C.2). This shows that the positive relationship between trading volume and $\hat{s}_{it}$ is not some spurious, in-sample relationship picked up when the multivariate linear probability model is estimated on the same sample as the panel regressions with trading volume and $\hat{s}_{it}$. The fact that the out-of-sample results are similar can also be anticipated from the fact that the (unreported) estimates for the linear probability model don't differ much across the two subperiods. The relationships between changes in characteristics and the propensity-to-sell seem to be quite stable over time. This can also be seen in Figure 2, which shows the univariate non-parametric estimates of this relationship, similar to Figure 1, but broken up into two subperiods. The solid lines represent the first subperiod (1985 – 1993), the dashed lines represent the second subperiod (1994 - 2002). The graphs show that for most characteristics, the relationship between characteristics changes and the propensity to sell is similar in both subperiods. Overall, the propensity to sell is lower in the second subperiod. For volatility, growth forecasts, and dividend yield, and perhaps also for the returns characteristics, the slopes are somewhat steeper in the second subperiod, but other than that there is not much difference across subperiods.

3.4 Further robustness checks

In additions to those reported in the tables, I have also explored regression specifications that include one-year-lagged volatility, lagged annual returns, absolute annual returns as an additional volatility measure, and lagged leverage, both with OLS and 2SLS. Chordia, Huh, and Subrahmanyam (2004) show that these variables have additional explanatory power for the cross-section of trading volume. I also find that they are significantly related to trading volume, in particular lagged returns and lagged volatility, but including these variables has little impact on the coefficient estimates for $\hat{s}_{it}$ and $\hat{s}_{it-1}$.

4. Concluding remarks

Style trading is important for the cross-section of trading volume. The results in this paper suggest that one key to explaining trading activity is to understand the forces that make investors specialize into diverse style categories and lead them to follow heterogeneous trading rules. At present neither the agency and industrial organization issues that might drive the specialization into styles, nor the notion of fixed trading rules as heuristics in investment decisions have a natural place in theoretical models of trading volume. Models that come closest are perhaps those where portfolios related to, say, size or value characteristics serve as hedging portfolios, such as in Lo and Wang (2001) and Mamaysky and Spiegel (2002), although it's not clear that this is the empirically most relevant mechanism. In a different literature, the emergence and evolution of investors' trading strategies has been studied in simulated stock markets with artificial agents that trade against each other. One of the findings in this literature is that certain heterogeneous trading strategies, reflecting heterogeneous beliefs, can emerge and co-exist (see, e.g., LeBaron 2000). While these papers give some hints as to what could perhaps drive investors' heterogeneity in trading styles, it seems that currently our theoretical understanding of this heterogeneity is still very limited.

The attempt in this paper to relate trading volume and style investing can be seen as part of a broader effort to use institutional and individual portfolio holdings data to understand why investors trade as, for example, in Grinblatt and Keloharju (2000, 2001), Barber and Odean (2001), Dorn and Huberman (2003), and Glaser and Weber (2004). A nice feature of the analysis here in this paper is that the connection of style-trading to changes in stock characteristics allows me to explain observed total trading volume, not only the trades in the portfolio holdings datasets. Given the apparent importance of heuristics for trading decisions, it seems that much can be learned from further investigations of how investors actually behave.

Appendix: Moving-Block Bootstrap

Let the data be a panel consisting of observations on the dependent variable, y_{it}, and a vector of explanatory variables, x_{it}, with $i = 1, \ldots, N$ and $t = 1, \ldots, T$. Consider a regression model

$$y_{it} = \alpha_t + \beta' x_{it} + \varepsilon_{it} \,, \tag{2}$$

where the residuals, ε_{it}, uncorrelated with x_{it}, can be dependent across time, t, and within each cross-section i. Consider the (OLS or 2SLS) estimator of β, denoted by β_T. Gonçalves and White (2004) show that under some weak conditions on the form of temporal dependence one can obtain consistent estimates of the standard errors of the linear regression coefficient estimates using a moving-block bootstrap (see, also, Künsch 1989; Hall, Horowitz and Jing 1995). To address both the temporal and the cross-sectional dependence in residuals, I draw blocks of entire cross-sections to construct the bootstrap samples. More precisely, the data can be organized into overlapping blocks of temporal length L, where block 1 includes the observations $\{x_{1t}, \ldots, x_{Nt}; y_{1t}, \ldots, y_{Nt}: t = 1, \ldots, L\}$, block 2 includes $\{x_{1t}, \ldots, x_{Nt}; y_{1t}, \ldots, y_{Nt}: t = 2, \ldots, L+1\}$, and so on. I then sample $K = T/L$ times from these T-L+1 overlapping blocks with replacement such that the resulting bootstrap sample, with the K sampled blocks concatenated in the order sampled, is of length T in the time dimension. The j-th bootstrap sample is then used to estimate the regression model (OLS or 2SLS) and record the coefficient estimates β_T^j. The resampling procedure and regression estimation is repeated 1,000 times. The variance-covariance of β_T^j across the bootstrap samples then provides an estimate of the variance-covariance of β_T.

The choice of block size is important to ensure that the variance-covariance estimate is accurate. Since in the setting here most of the autocorrelation comes from the fact that the quarterly cross-sections are overlapping annual observation intervals, this suggests that the block size should be four quarters at a minimum. To allow for further autocorrelation, I use 12 quarters (T = 72 quarters for the full sample).

References

Admati, A.R., and P. Pfleiderer, 1998, A Theory of Intraday Patterns: Volume and Price Variability, *Review of Financial Studies* 1, 3-40.

Almazan, Andres, Keith C. Brown, Murray Carlson, and David A. Chapman, 2004, Why Constrain your Mutual Fund Manager?, *Journal of Financial Economics*, forthcoming.

Atkins, A., and E. Dyl, 1997, Market Structure and Reported Trading Volume, *Journal of Financial Research* 20, 291-304.

Barber, B.M, and T. Odean, 2001, Boys Will be Boys: Gender, Overconfidence, and Common Stock Investment, *Quarterly Journal of Economics* 116, 261-292.

Barberis, N., and A. Shleifer, 2003, Style Investing, *Journal of Financial Economics* 68, 161-199.

Black, F., 1986, Noise, *Journal of Finance* 56, 529-543.

Brown, S.J., and W.N. Goetzmann, 1997, Mutual Fund Styles, *Journal of Financial Economics* 43, 373-399.

Chan, L., H. Chen, and J. Lakonishok, 2002, On Mutual Fund Investment Styles, *Review of Financial Studies* 15, 1407-1437.

Chordia, T., S.-W. Huh, and A. Subrahmanyam, 2004, The Cross-Section of Expected Trading Activity, Working paper, UCLA.

Chu, J., and C.L. and Osler, 2004, Microstructual Benefits of Imperfectly Rational Trading, Working Paper, Brandeis University.

Cremers, K.J.M., and J. Mei, 2004, Turning Over Turnover, Working Paper, Yale School of Management and NYU.

Dorn, D. and G. Huberman, 2003, Talk and Action: What Individual Investors Say and What They Do, Working Paper, Columbia University.

Dow, J., and G. Gorton, 1997, Noise Trading, Delegated Portfolio Management, and Economic Welfare, *Journal of Political Economy* 105, 1024-1050.

Diether, K.B, C.J. Malloy, A. Scherbina, 2002, Differences in Opinion and the Cross-Section of Stock Returns, *Journal of Finance* 57, 2113-2141.

Fung, W., and D.A. Hsieh, 1997, Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds, *Review of Financial Studies* 10, 275-302.

Glaser, M., and M. Weber, 2004, Overconfidence and Trading Volume, Working Paper, University of Mannheim.

Gonçalves, S., and H. White, 2004, Bootstrap Standard Error Estimation for Linear Regressions, Working Paper, UC San Diego.

Grinblatt, M., and M. Keloharju, 2001, What Makes Investors Trade? *Journal of Finance* 66, 589-616.

Harris, M., and A. Raviv, 1993, Differences in Opinion Make a Horse Race, *Review of Financial Studies* 6, 473-506.

He, H., and J. Wang, 1995, Differential Information and Dynamic Behavior of Stock Trading Volume, *Review of Financial Studies* 8, 919-972.

Hong, H., and J.C. Stein, 2003, Simple Forecasts and Paradigm Shifts, Working Paper, Princeton University and Harvard University.

Hotchkiss, E.S., and S. Lawrence, 2003, Empirical Evidence on the Existence of Dividend Clienteles, Working Paper, Boston College.

Hall, P., J. L. Horowitz, and B. Jing, 1995, On Blocking Rules for the Bootstrap with Dependent Data, *Biometrika* 82, 561-74.

Huddart, S., M. Lang, and M. Yetman, 2003, Psychological Factors, Stock Price Paths and Trading Volume, Working Paper, Pennsylvania State University.

Kandel, E., and N.D. Pearson, 1995, Differential Interpretation of Public Signals and Trade in Speculative Markets, *Journal of Political Economy* 103, 831-872.

Karpoff, J.M., 1987, The Relation between Price Changes and Trading Volume: A Survey, *Journal of Financial and Quantitative Analysis* 22, 109-126.

Künsch, H.R., 1989, The Jackknife and the Bootstrap for General Stationary Observations, *Annals of Statistics* 17, 1217-41.

LeBaron, B., 2000, Agent Based Computational Finance: Suggested Readings and Early Research, *Journal of Economic Dynamics and Control* 24, 679-702.

Lo, A.W., and J. Wang, 2000, Trading Volume: Definitions, Data Analysis, and Implications of Portfolio Theory, *Review of Financial Studies* 13, 267-300.

Lo, A.W., and J. Wang, 2001, Trading Volume: Implications of an Intertemporal Capital Asset Pricing Model, Working Paper, MIT.

Mamaysky, H., and M. Spiegel, 2002, A Theory of Mutual Funds: Optimal Fund Objectives and Industrial Organization, Working Paper, Yale School of Management.

Massa, M., 2000, Why So Many Mutual Funds? Mutual Funds, Market Segmentation, and Financial Performance, Working Paper, INSEAD.

Mullainathan, S., 2002, Thinking Through Categories, Working Paper, MIT.

Ross, S.A., 1989, Discussion: Intertemporal Asset Pricing, in: *Theory of Valuation: Frontiers of Modern Financial Theory*, edited by S. Bhattacharya and G.M. Constantinides, Totowa, N.J., Rowman & Littlefield.

Varian, H.R., 1989, Differences in Opinion in Financial Markets, in C.C. Stone (Ed.) *Financial Risk: Theory, Evidence and Implications*, *Proceedings of the Eleventh Annual Economic Policy Conference of the Federal Reserve Bank of St. Louis*, Kluwer Academic, Norwell, MA.

Wang, J., 1994, A Model of Competitive Stock Trading Volume, *Journal of Political Economy* 102, 127-168.

Wermers, R., 1999, Mutual Fund Herding and the Impact on Stock Prices, *Journal of Finance* 54, 581-622.

Wermers, R., 2002, A Matter of Style: The Causes and Consequences of Style Drift, Working Paper, University of Maryland.

Table 1: Summary Statistics

This table shows summary statistics using stocks that have a valid observation on each of the 12 stock characteristics in a given quarter. The sample period runs from the first quarter of 1984 to the last quarter of 2002. Stock characteristics are first transformed into percentile ranks ranging from 0 to 99. For stock-level characteristics the unit of observation is a stock. For fund-level statistics, the unit of observation is a fund. Mean fund characteristics are calculated as a value-weighted (using portfolio weights) average of the characteristics of the stocks in a fund's portfolio each quarter. The first block of rows shows statistics on the cross-sectional distribution of these fund mean characteristics. The second block shows the sample correlation matrices for stocks' characteristics (upper triangular part) and fund mean characteristics (lower triangular part). Medians, percentiles, and correlations are calculated cross-sectionally each quarter and are then averaged over time.

	Return(-12:-1)	Return(-3:-1)	Return(-36:-1)	Sales/MarketCap	Earnings/Assets	Liabilities/MarketCap	Cash/CurrLiab	LTGrowthForecast	DividendYield	Beta	Volatility	MarketCap
	Cross-sectional distribution of mean characterics of fund portfolios											
Mean	58.35	54.42	62.06	35.93	57.91	42.86	49.00	46.46	59.54	49.58	35.20	83.50
Median	58.91	54.81	63.11	35.08	59.63	42.26	47.25	44.64	61.66	48.79	31.62	88.82
5^{th} pct.	36.44	34.59	37.68	13.44	31.90	17.24	30.21	18.44	29.99	24.70	15.21	51.20
95^{th} pct.	78.11	72.80	82.16	60.51	78.18	71.59	72.77	78.80	85.30	73.82	65.30	97.16
					-- correlation among stock characteristics --							
Return(-12:-1)		0.49	0.55	-0.27	0.05	-0.23	-0.04	0.00	0.10	-0.09	-0.22	0.29
Return(-3:-1)	0.51		0.28	-0.15	0.03	-0.13	-0.01	-0.03	0.06	-0.03	-0.09	0.16
Return(-36:-1)	0.61	0.29		-0.43	0.30	-0.39	0.03	0.16	0.07	-0.01	-0.23	0.36
Sales/MarketCap	-0.25	-0.12	-0.47		-0.31	0.74	-0.42	-0.39	0.13	-0.12	0.00	-0.30
Earnings/Assets	0.09	0.04	0.38	-0.46		-0.51	0.19	0.10	0.08	-0.01	-0.18	0.21
Liabilities/MarketCap	-0.23	-0.12	-0.47	0.82	-0.62		-0.51	-0.51	0.26	-0.20	-0.14	-0.11
Cash/CurrLiab	0.00	0.00	0.15	-0.53	0.23	-0.65		0.34	-0.30	0.20	0.27	-0.11
LTGrowthForecast	0.16	0.06	0.40	-0.53	0.27	-0.70	0.60		-0.60	0.41	0.50	-0.23
DividendYield	-0.07	-0.03	-0.20	0.36	-0.07	0.55	-0.63	-0.83		-0.39	-0.67	0.42
Beta	0.03	0.04	0.17	-0.18	0.07	-0.37	0.41	0.67	-0.66		0.44	-0.07
Volatility	-0.04	-0.01	0.02	-0.20	-0.13	-0.37	0.59	0.68	-0.85	0.64		-0.56
MarketCap	0.14	0.07	0.19	-0.17	0.24	0.01	-0.30	-0.34	0.55	-0.25	-0.65	
	-- correlation among mean fund characteristics --											

average number of stocks:	1730	average number of stocks per fund:	39	(mean)
average number of funds:	3788		25	(median)

Table 2: Cross-Sectional Distribution of Trading Volume (Turnover)

The sample period runs from the first quarter of 1985 to the fourth quarter of 2002. NYSE/AMEX and Nasdaq turnover is computed monthly as the number of shares traded during a month divided by the average of the number of shares outstanding at the beginning and at the end of the month. This monthly turnover is then summed over 12-month periods to an annual turnover figure. Mutual funds selling volume is calculated for each stock from the mutual fund portfolio holdings data as the number of shares (split-adjusted) of this stock sold by mutual funds during an annual period, divided by the number of shares held by these funds at the beginning of the annual period. The statistics shown in the table are first calculated for each (overlapping) quarterly cross-section and are then averaged over time.

	Mean	Median	Std.dev.	5th pctile	95th pctile
NYSE and AMEX	87%	70%	65%	23%	203%
Nasdaq	174%	127%	154%	28%	476%
Mutual funds selling volume (from quarterly holdings)	42%	39%	24%	6%	87%

Table 3: Changes in Stock Characteristics and Mutual Funds' Propensity to Sell

The first column reports the mean absolute change in stock characteristics, where the change is calculated for each stock in each quarter over an annual period. Stock characteristics are first transformed into percentile ranks, ranging from 0 to 99. The mean absolute change is first calculated within each overlapping quarterly cross-section, and is then averaged across time. The second column shows point estimates of the slope coefficients in the linear probability model, estimated as a univariate model with the absolute change of only one characteristic as the explanatory variable. The third column shows point estimates for the slope coefficients in the multivariate version, where the absolute changes of all characteristics jointly serve as explanatory variables. The dependent variable is an indicator variable, s_{ijt}, that equals one if fund j sold all its holdings of stock i during period t, and zero otherwise. The models are estimated first cross-sectionally each quarter across all fund-stock observations, and then the coefficient estimates are averaged over time. Standard errors of the coefficient estimates (not reported) are negligibly small.

		Change in propensity to sell for a one-standard deviation change in characteristcs	
Characteristic	Mean absolute change per year	Univariate slopes (in %)	Multivariate slopes (in %)
Return(-12:-1)	38.78	5.83	2.68
Return(-3:-1)	39.03	3.94	1.25
Return(-36:-1)	25.12	5.83	1.94
Sales/MarketCap	13.95	7.60	3.30
Earnings/Assets	23.65	1.85	1.09
Liabilities/MarketCap	13.21	7.30	0.73
Cash/CurrLiab	17.65	1.80	0.30
LTGrowthForecast	16.45	5.17	2.19
DividendYield	9.42	5.01	1.14
Beta	19.25	3.39	1.03
Volatility	15.73	3.80	1.67
MarketCap	8.26	6.51	3.09

Table 4: Explaining the Cross-Section of Trading Volume, OLS Regressions

This table reports results of OLS panel regressions of annual trading volume from CRSP on the propensity to sell, $\hat{s}_{it}$ and $\hat{s}_{it-1}$, calculated as the fitted values from the multivariate linear probability model. For NYSE/AMEX (Panel A) and Nasdaq stocks (Panel B), trading volume is calculated as turnover, i.e., the monthly number of shares traded, divided by the average of the beginning and end-of-month number of shares outstanding, and summed up over a 12-month period. Mutual fund trading volume in Panel B is calculated for each stock as the number of shares (split-adjusted) of this stock sold by mutual funds during an annual period, divided by the number of shares held by these funds at the beginning of the annual period. Volatility is calculated as the squared daily return over an annual period. Size is the four-quarter lagged market capitalization, and Price is the four-quarter lagged stock price (i.e., both measured at the beginning of the annual period t). Forecast dispersion is a 12-month average of the Diether, Malloy, and Scherbina (2002) analyst forecast dispersion measure. All explanatory variables, including $\hat{s}_{it}$ and $\hat{s}_{it-1}$, are demeaned and standardized to unit cross-sectional standard deviation within each quarterly cross-section. Standard errors shown in parentheses are obtained from a moving-block bootstrap, described in the appendix, which accounts for both cross-sectional and temporal dependence in residuals. The sample period runs from the first quarter of 1985 to the fourth quarter of 2002.

	Propensity to sell from fitted linear probability model		Control variables				
	$\hat{s}_{it}$	$\hat{s}_{it-1}$	Volatility	Size	Price	Forecast Dispersion	adj. R^2
Panel A: Total trading volume (turnover), NYSE and AMEX stocks							
A.1	0.18	0.08					10.2%
	(0.02)	(0.02)					(0.02)
A.2	0.14	0.10	0.27	0.08	0.19		25.2%
	(0.01)	(0.01)	(0.04)	(0.01)	(0.04)		(0.03)
A.3	0.12	0.09	0.35	0.05	0.19	0.02	29.1%
	(0.01)	(0.01)	(0.07)	(0.01)	(0.04)	(0.01)	(0.04)
Panel B: Total trading volume (turnover), Nasdaq stocks							
B.1	0.35	0.17					5.5%
	(0.04)	(0.02)					(0.01)
B.2	0.32	0.25	0.64	0.60	0.42		30.3%
	(0.03)	(0.03)	(0.15)	(0.05)	(0.13)		(0.05)
B.3	0.26	0.26	1.03	0.52	0.50	0.04	35.0%
	(0.04)	(0.03)	(0.16)	(0.05)	(0.14)	(0.03)	(0.05)
Panel C: Mutual fund selling volume (turnover, from quarterly holdings data), NYSE, AMEX, and Nasdaq stocks							
C.1	0.06	0.02					9.2%
	(0.00)	(0.00)					(0.01)
C.2	0.06	0.02	0.06	0.04	0.03		16.1%
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)		(0.02)
C.3	0.06	0.02	0.07	0.02	0.03	0.00	16.1%
	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.00)	(0.02)

Table 5: Explaining the Cross-Section of Trading Volume, 2SLS Regressions

The regressions presented in this table are similar to those of Table 4, but now estimation is done with 2SLS. The explanatory variables $\hat{s}_{it}$ and $\hat{s}_{it-1}$ are instrumented with $\hat{s}_{it}$* and $\hat{s}_{it-1}$*, which are calculated in exactly the same way as fitted values of the multivariate linear probability model, but using only characteristics that are not based on stock price (Earnings/Assets, Liabilities/Assets, Cash/CurrLiab, and one-year sales growth). All explanatory variables, including $\hat{s}_{it}$ and $\hat{s}_{it-1}$ and the instruments, are demeaned and standardized to unit cross-sectional standard deviation within each quarterly cross-section. Standard errors shown in parentheses are obtained from a moving-block bootstrap, described in the appendix, which accounts for both cross-sectional and temporal dependence in residuals. The sample period runs from the first quarter of 1985 to the fourth quarter of 2002.

	Propensity to sell from fitted linear probability model		Control Variables			Instruments	
	$\hat{s}_{it}$	$\hat{s}_{it-1}$	Volatility	Size	Price	$\hat{s}_{it}$*	$\hat{s}_{it-1}$*
Panel A: Total trading volume (turnover), NYSE and AMEX stocks							
First Stage:	X		0.31	-0.18	0.02	0.20	0.07
			(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
First Stage:		X	0.18	-0.16	-0.10	0.12	0.24
			(0.03)	(0.01)	(0.03)	(0.01)	(0.01)
Second stage:	0.17	0.16	0.24	0.10	0.20		
	(0.04)	(0.03)	(0.03)	(0.01)	(0.04)		
Panel B: Total trading volume (turnover), Nasdaq stocks							
First Stage:	X		0.18	-0.02	-0.02	0.21	0.07
			(0.04)	(0.02)	(0.03)	(0.01)	(0.01)
First Stage:		X	0.10	-0.04	-0.09	0.13	0.24
			(0.02)	(0.02)	(0.03)	(0.01)	(0.02)
Second stage:	0.71	0.49	0.52	0.62	0.47		
	(0.11)	(0.06)	(0.12)	(0.05)	(0.12)		
Panel C: Mutual fund selling volume (turnover), NYSE, AMEX, and Nasdaq stocks							
First Stage:	X		0.27	-0.15	-0.01	0.21	0.07
			(0.02)	(0.02)	(0.02)	(0.01)	(0.00)
First Stage:		X	0.16	-0.15	-0.10	0.13	0.24
			(0.02)	(0.01)	(0.02)	(0.01)	(0.01)
Second stage:	0.06	0.03	0.06	0.04	0.03		
	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)		

Table 6: Subperiod and Out-of-Sample Tests, 2SLS (second-stage) regressions

The 2SLS regressions presented in this table are similar to those in Table 5 (only the second stage is shown here), but they run here over the subsamples 1985-1993 and 1994-2002. For the second subperiod, the table also shows out-of-sample regressions, which means that the multivariate linear probability model is estimated over the first subperiod. The estimated model is then used to generate the fitted propensity to sell $\hat{s}_{it}$ and $\hat{s}_{it-1}$ and the instruments $\hat{s}_{it}$* and $\hat{s}_{it-1}$* in the second subperiod. Standard errors shown in parentheses are obtained from a moving block-bootstrap, described in the appendix, which accounts for both cross-sectional and temporal dependence in residuals.

	Propensity to sell from fitted linear probability model		Control Variables		
	$\hat{s}_{it}$	$\hat{s}_{it-1}$	Volatility	Size	Price
Panel A: Total trading volume (turnover), NYSE and AMEX stocks					
A.1 (1985 - 1993)	0.11	0.13	0.15	0.10	0.09
	(0.04)	(0.03)	(0.01)	(0.02)	(0.02)
A.2 (1994 - 2002)	0.22	0.17	0.31	0.09	0.27
	(0.03)	(0.05)	(0.02)	(0.01)	(0.05)
A.3 (1994 - 2002, out-of-sample)	0.23	0.18	0.30	0.09	0.27
	(0.03)	(0.05)	(0.02)	(0.01)	(0.05)
Panel B: Total trading volume (turnover), Nasdaq stocks					
B.1 (1985 - 1993)	0.35	0.47	0.10	0.43	0.01
	(0.17)	(0.08)	(0.02)	(0.03)	(0.06)
B.2 (1994 - 2002)	0.84	0.46	0.63	0.70	0.61
	(0.06)	(0.07)	(0.05)	(0.02)	(0.11)
B.3 (1994 - 2002, out-of-sample)	0.81	0.46	0.64	0.71	0.61
	(0.07)	(0.07)	(0.05)	(0.02)	(0.11)
Panel C: Mutual fund selling volume (turnover), NYSE, AMEX, and Nasdaq stocks					
C.1 (1985 - 1993)	0.05	0.02	0.07	0.05	0.02
	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
C.2 (1994 - 2002)	0.06	0.03	0.06	0.03	0.04
	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)
C.3 (1994 - 2002, out-of-sample)	0.06	0.04	0.06	0.03	0.04
	(0.00)	(0.00)	(0.02)	(0.02)	(0.01)

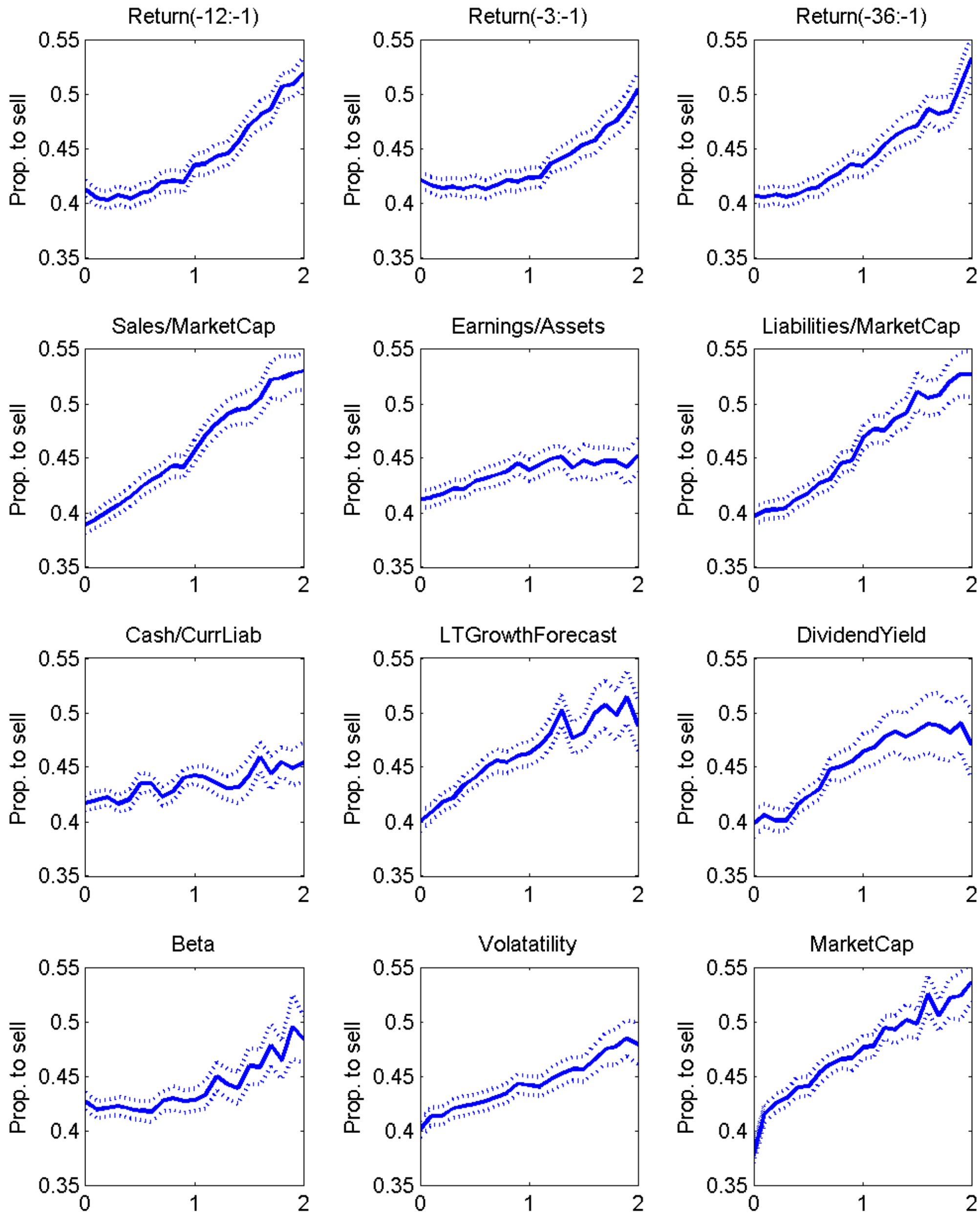

Figure 1. Univariate non-parametric estimates of the propensity to sell as a function of absolute changes in stock characteristics. In each cross-section t, and separately for each stock characteristic, stocks are binned by the absolute change in the characteristic, standardized by the cross-sectional standard deviation of the characteristic. The width of each bin is 0.1 standard deviations. For each bin, the mean of s_{ijt}, the selling indicator variable, is calculated. For each characteristic, the solid lines in the figure show, for each bin from 0 to 2 standard deviations (i.e., bin 1 to bin 20), the time-series average of these means. The dotted lines represent two-standard error bands.

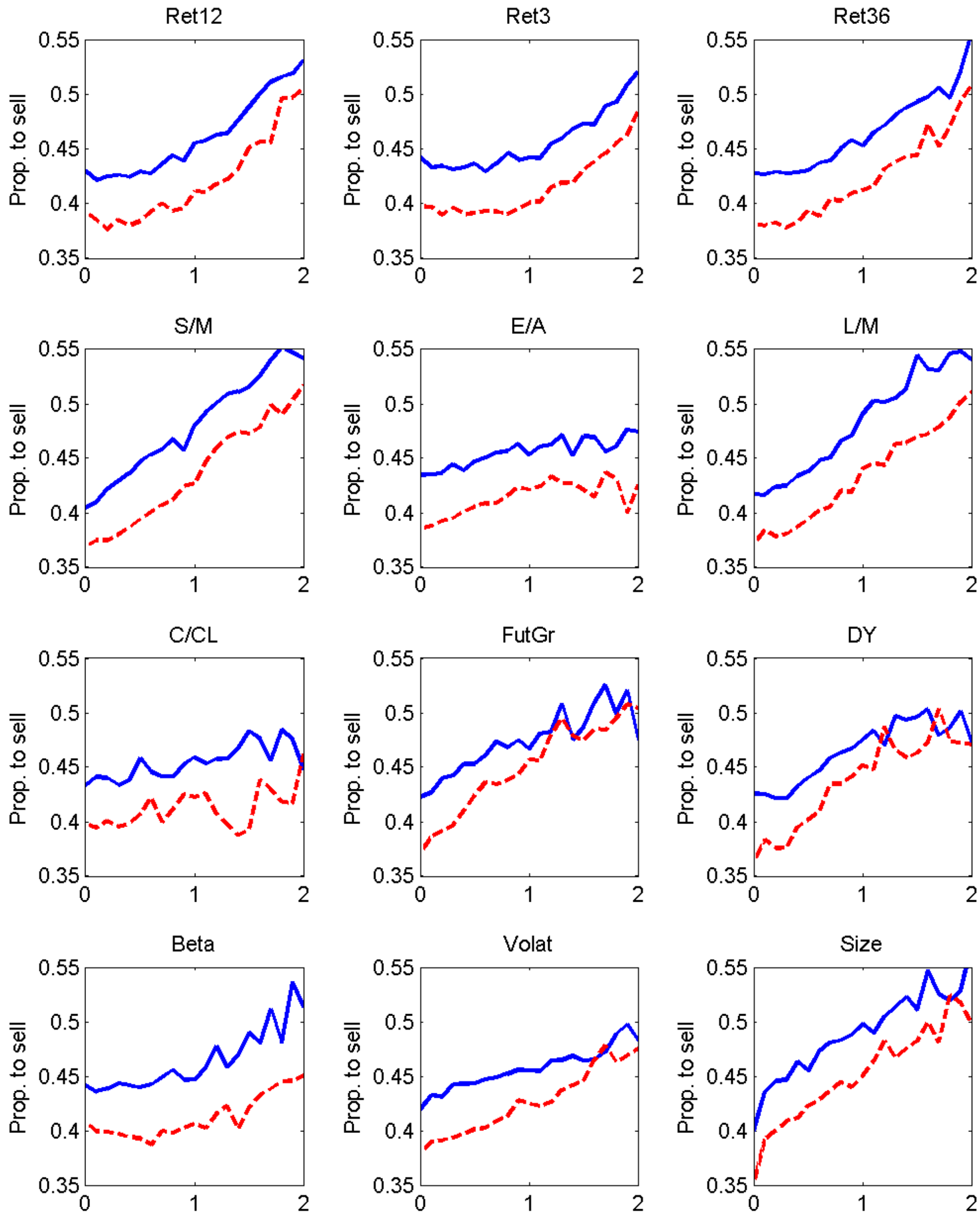

Figure 2. Univariate non-parametric estimates of the propensity to sell as a function of absolute changes in stock characteristics for subperiod samples. The figures are identical to those in Figure 1, except for the different sample periods. Results for 1985 - 1993 are shown as solid lines, results for 1994 – 2002 are shown as dashed lines.