Factor Crowding and Liquidity Exhaustion

Joseph M. Marks*

December 22, 2016

Abstract

Well-known anomalies and stable patterns in equity returns are widely employed to guide active stock selection. The use of overlapping multifactor models built on these patterns induces correlated trade across investors. Consistent with correlated trade, a stock with a strong signal from a parsimonious stock selection model is associated with greater future trade activity and lower return volatility. Stocks favored by the model also experience significant decreases to their level of liquidity and increases in the degree to which their liquidities covary. These results suggest that correlated trading among institutions is an important source of commonality in liquidity, and that measures of portfolio liquidity that ignore these changes understate risk.

*Assistant Professor of Finance at Bentley University, jmarks@bentley.edu, (781) 891-2654, 175 Forest St. AAC 273, Waltham, MA, 02452. I am grateful to Chenguang Shang, David Simon, Jim Musumeci, Jay Sultan, and participants at the Bentley Finance Research seminar during the Spring of 2016 for helpful feedback and comments. All errors remain my own.

1 Introduction

Stock liquidity has long been a topic of interest and research. Over the past thirty years or so, this body of literature has evolved significantly as the focus of research develops and shifts. Early articles concerning stock liquidity focused on cross-sectional variation and its determinants. Amihud and Mendelson (1986), Glosten and Harris (1988), and Stoll (1989) measure and decompose the bid-ask spread with a focus on the component due to information and adverse selection. Researchers have also used transaction-level data to characterize liquidity in terms of the price response to order flow, among them Brennan and Subrahmanyam (1996), Hasbrouck (1991), and Easley and O'Hara (1987). Amihud (2002) proposes a widely-used measure of liquidity that is highly correlated to computationally intensive variables that require transaction-level data, yet is much simpler to calculate because it employs daily data. Beginning with Chordia, Roll, and Subrahmanyam (2000) and Hasbrouck and Seppi (2001), the focus shifted from the level of individual stock liquidity to documenting a systematic component to changes in liquidity. The finding that a systematic liquidity factor explains a significant portion of the time-variation in liquidity motivates the studies of Pastor and Stambaugh (2003) and Acharya and Pedersen (2005), both of which find a stock's sensitivity to a systematic liquidity factor helps explain cross-sectional heterogeneity in returns, and hence a liquidity premium in the sense of classical asset pricing models.

This paper extends previous work in this field in order to better understand factors that induce common variation in stock liquidity. While it is well-established that the liquidities of individual stocks do covary, the reasons for this commonality are less well understood. One branch of literature focuses on the role that market makers play in determining liquidity through managing inventory risk in the presence of external funding constraints. Grossman and Miller (1988) and Brunnermeier and Pedersen (2009) are among the articles providing evidence that market makers do affect time variation in liquidity through these channels. Lo and Wang (2000) take a different approach and demonstrate that if returns obey a linear factor structure, then under suitable conditions trading volume will follow the same structure. Thus, the state variables represented by the factors determine both covariance in returns and covariance in trading volume and presumably liquidity comovement as well. Kamara, Lou, and Sadka (2008) study changes in liquidity commonality and find that over time comovement of the liquidities of large stocks has increased while comovement of small stock liquidity has decreased. They attribute this result to the large growth in institutional ownership that would affect stock liquidity as institutions tend to trade the stocks together in a correlated manner as is the case, for example, for index funds.

This paper builds upon these earlier studies to further demonstrate that correlated trading affects both the level of stock liquidity and comovement of liquidity. To capture correlated trading, a multifactor model based on asset pricing anomalies is used to generate the types of signals that active portfolio managers employ when selecting stocks. I find that a strong signal to buy or sell a stock is associated with greater trade activity, lower future volatility consistent with a directional bias in trade, a lower level of stock-specific liquidity, and greater comovement of individual stock liquidity within the set of stocks favored by the multifactor model. This is an extension of the result in Kamara, Lou, and Sadka (2008) which ties correlated trading more broadly to firm size based on the notion that large institutional investors cannot feasibly trade small, illiquid stocks. The multifactor model used in this work provides more specific signals of the individual stocks that many investors may trade.

The finding that correlated trading affects liquidity commonality has important implications for measures of portfolio risk. Following dramatic market events such as the default of Long-Term Capital Management, the Quant Crisis of August 2007, and the 2008 Financial Crisis, there is an increased awareness that overlapping portfolio positions may introduce significant risk. The rapid liquidation of one portfolio can cause a severe market disruption that adversely affects investors holding similar positions. If a loss is large enough, it could trigger a feedback cycle in which other investors reduce their positions to manage risk and leverage, and in so doing amplify the original disruption. Although the evidence is inconclusive, this is a commonly cited cause of the perverse behavior of widely-used stock selection factors in August 2007. In response to this episode, many investors began asking if these stock selection factors were crowded and could no longer be relied upon to deliver the same performance as they had in the past. Aside from such dramatic events, this paper demonstrates that the regular correlated trading of many investors does affect stock-level liquidity. If measures of portfolio liquidity are based on estimates extracted from return data prior to the stock's entry into the portfolio, liquidity risk may be understated to the extent that the increase in liquidity commonality is ignored.

A focus on the liquidity dimension associated with factor crowding is a novel feature of this study. Previous studies have concentrated on the profitably of potentially crowded factors and the primary question of interest has been if active investors have driven the expected profits to zero, much as increasingly aggressive arbitrageurs would be expected to correct mispricing through their exploitation. Examples of these studies include Gustafson and Halper (2010) and Cahan and Luo (2013), and the evidence they provide suggests that factors are not fully crowded in this sense. Further evidence comes from Verbeek and Wang (2013), who demonstrate that using the SEC-mandated quarterly disclosures to mimic the holdings of active mutuals funds provides the same performance as the target funds, and hence the strategies appear to have excess capacity. Rather than a focus on profitiability, the present work is focused on how the liquidities of individual stocks may be affected by active institutional investors with overlapping models and positions. There is good reason to expect that institutional trading would affect stock liquidity. An obvious example of this effect is the large return documented of stocks upon their inclusion in a major index. Early studies in this area by Harris and Gurel (1986) and Shleifer (1986) concluded that a significant portion of this return is temporary and subsequent research has supported this conclusion. In the case of index inclusion, passive index portfolio managers acquire the stock to replicate the index, and though the affected stocks are highly liquid large cap stocks and demand is clearly uninformed, the limited liquidity of the market cannot absorb the demand shock without a price increase. In a similar vein, if many investors use similar models to identify stocks for active portfolios, then they may trade in a correlated manner that exhausts the liquidity available in the stocks that they trade.

The remainder of this paper is organized as follows. Section 2 proposes a simple multifactor model for stock selection, provides estimates of its parameters, and documents its out-of-sample performance. Section 3 contains the main results concerning the effects of correlated trading on the level of trade activity, the level of stock volatility, the level of stock liquidity, and changes in liquidity commonality. Section 4 provides concluding remarks.

2 Multifactor Models

Active investors employ a variety of quantitative approaches to stock selection. One prominent example is technical analysis, which in its simple form can employ the type of pattern recognition usually associated with it, and in its more modern manifestation can be based on signals generated by advanced time series models such as the fractional vector autoregression with error correction model of Caporin, Ranaldo, and de Magistris (2013). Instead of restricting their input data to historical prices, other investors adopt a more comprehensive approach that also relies on fundamental data. A simple fundamental model would calculate ratios obtained from financial statements and then model returns as a function of these ratios in a regression framework. Other fundamental approaches depart from the familiar regression framework but retain the focus on financial ratios, for instance generalized data envelopment analysis developed in Edirisinghe and Zhang (2007). Still other investors expand their data to include nearly any source that can somehow be quantified, and textual analysis of verbal sources such as corporate disclosures and management interviews is a prime example of how this may be done.[1]

While there are many approaches to active stock selection, they are not equally popular. A great many active investors employ linear multifactor models during the investment process, and some evidence of their popularity can be inferred from their commercial success. The most obvious way of using this type of model is to generate expected return signals to aid in stock selection. Indeed, even a manager with a traditional focus on fundamental research may subscribe to a quantitative multifactor model as a check or additional signal to inform their stock choices. Even if the model is not used directly for stock selection, it is difficult to ignore measures of active portfolio risk captured by the factors on which these models are based. Because performance attribution is often conducted within this framework, portfolio managers that ignore the factors altogether may end up explaining large bets that they did not intend to take. Commercial multifactor linear models also

[1] With the benefit of computing power, the field of textual analysis has rapidly developed during the past ten years. Werner and Frank (2004) study 45 stocks in the Dow Jones Industrial Average and Dow Jones Internet Commerce indexes and find a significant relation between the tone of messages posted on two popular internet sites and stock volatility. Tetlock (2007) uses a column from the *Wall Street Journal* as the textual source and finds short term predictability in returns and volume based on the tone of the column. Feldman, Govindaraj, Livnat, and Segal (2010) study the management discussion and analysis sections of quarterly and annual financial statements and find that tone changes have incremental predictive power relative to earnings surprises and measurements of accruals. McKay Price, Doran, Peterson, and Bliss (2012) study transcripts of public conference calls with management and, similar to Tetlock (2007), find that linguistic tone predicts abnormal returns and trading volume.

have innate appeal because of their similarity to (and under certain conditions, consistency with) arbitrage pricing theory. For these reasons, multifactor fundamental models enjoy widespread use.

The MSCI Barra USE3 and Standard and Poor's Capital IQ models are two of the most popular models within the institutional investment community. Institutional investors own and control a large fraction of equity markets, so their use of these models creates the opportunity for the models to influence the behavior of the stocks they trade. Both the models marketed by Barra and S&P Capital IQ are based on a large number of firm-specific variables that past research has shown either predict stock returns or are related to risk. Many of these variables are gleaned from the anomalies literature, including familiar factors capturing firm size, book-to-price, momentum, volatility, and liquidity. In the case of the USE3 model, there are 39 variables (so called *descriptors*) organized into 12 groups referred to as risk indexes.[2] For example, the growth risk index is one way the model attempts to capture the return spread between value and growth stocks, and it relies on historical dividend payout ratio, growth rate in total assets, five-year trailing earnings growth rate, analyst-predicted growth rate, recent earnings changes, and variability in capital structure.

The question this paper addresses is whether or not correlated trading among investors affects the liquidities of the stocks they trade. Given the popularity of multifactor models, it is reasonable to use one as a basis for measuring the stocks that are likely to be traded by active institutional managers. However, proposing a specific model creates a tension between two considerations. On one hand, a more detailed model with a greater number of factors is a more realistic representation of the models actually employed by investors, and will therefore better capture the trade activity of some investors. Yet, greater model specificity is also more likely to diverge from the consensus that gives rise to correlated trading. For instance, many investors in some manner use historical standard deviation or market beta as a measure of volatility; Aside from the actual subscribers to USE3, fewer investors use other descriptors included in the volatility risk index, such as serial correlation, implied volatility from options, the range between 52-week high and low prices, and the product of beta and standard deviation. Many active managers pride themselves on their development of unique factors and proprietary models, while many more do not possess the resources necessary to develop the models in-house.

[2]While there are 12 risk indexes, two of them are non-linearities in other risk indexes (firm size and beta) and hence they are not separate, independent factors.

To balance model specificity with the desire to capture general patterns in trade, the multifactor stock selection model that I employ in this study uses one descriptor from each risk index in the USE3 model. Table 1 provides a summary of the included variables, all of which are commonly seen in empirical studies focused on the cross-section of stock returns. There is a vast literature on stock anomalies and many other reasonable variables could be chosen in place of or in addition to those listed in Table 1. However, these particular factors are especially relevant because they constitute the core of popular, commerically available quantitative models. To ensure that the empirical results that appear later in this paper are not tied to this specific model, I have repeated the following analyses using an alternative model after choosing reasonable substitutes for each descriptor. The results obtained using the alternative model are similar to those presented here. The essential characteristic shared by these models is that they capture common investment signals shown to explain cross-sectional variation in returns in a parsimonious manner.

In addition to specifying the stock selection model used to generate trade signals, an important empirical decision must be made concerning the sample period. The objective is to capture a time interval during which active investors could combine many investment signals to generate stock rankings for a large universe. Doing so requires the availability of the many types of data used to calculate the selection factors, and the technological innovations that have made it possible to access and store this data and estimate the model in a production environment. These conditions would not have been satisfied, for instance, during the 1970s, when accessing financial data and statistical computing were in their infancy compared to the present day. Jacobs and Levy (1988) proposes and tests one the earliest comprehensive multifactor models of this type, and as the authors were institutional portfolio managers at the time of publication, could be considered the beginning of this period. While it was feasible for more sophisticated investors to employ this approach as early as the late 1980s, it was not until somewhat later that many other investors became informed of this approach and overcame the barriers to entry. These barriers have been steadily declining as the development of front-end user interfaces make it increasingly simple to access multiple databases, specify a model, and estimate its parameters. With these considerations in mind, I begin the sample period in January 1995 and extend through December 2014.

2.1 Model Estimation and In-Sample Performance

The variables listed in Table 1, hereafter referred to as factors, are used in a cross-sectional regression to estimate return premia. The estimated premia are then used to generate out-of-sample expected returns. In particular, for each month t in the sample period, the following regression is estimated:

$$r_{it} = \alpha_t + \sum_{k=1}^{9} \lambda_t^k z_{it-1}^k + \epsilon_{it} \tag{1}$$

where r_{it} is the monthly total return on stock i earned from the end of month $t-1$ through the end of month t, z_{it-1}^k is stock i's standardized exposure to factor k at the end of month $t-1$, and λ_t^k is the premium associated with factor k during month t. Equation (1) is estimated using both ordinary and weighted least squares, the latter weighting observations by the inverse of the residual standard deviation obtained from estimating the Fama-French three-factor model during the prior year with daily data. Each month, standardized exposures are produced for each stock for each factor by subtracting the cross-sectional average of the variable and then dividing the difference by the cross-sectional standard deviation. As with the factors themselves, the approach to variable measurement and estimation employed here is standard and is described in detail by Chincarini and Kim (2006) and MSCI Barra.[3]

Table 2 presents summary statistics for the variables used to calculate factor exposures, monthly returns, and liquidity level. The Amihud (2002) measure of liquidity captures the notion of price impact by calculating the ratio of the absolute daily return to dollar volume traded:

$$AL_{is} = \frac{|r_{is}|}{v_{is} p_{is}} \tag{2}$$

where r_{is}, v_{is}, and p_{is} are daily observations of return, volume, and closing price. The statistics in the table are time-series means of cross-sectional means. The time-averaging of Amihud liquidity obscures the strong negative trend in the cross-sectional average during this period, and in most of the following analyses the focus will be on first differences in monthly averages. The statistics in Table 2 provide basic measures of the factor data before it is standardized.

As indicated in Table 1, the Barra USE3 model includes a volatility risk index as is typical

[3] See Chapter 5 of Chincarini and Kim (2006) and Section 5 of the USE3 Handbook detailing factor exposure calculation and model estimation.

for a multifactor model of this type. In Section 3, one area of focus is the change in stock-level volatility following a signal generated by using (1) to forecast expected returns. In these cases, the inclusion of volatility as one of the factors would be inappropriate, especially if the estimated volatility premium is negative.[4] Even in cases when the focus is on liquidity and not volatility, given a relation between these two variables, it would be confounding to include a volatility factor in the stock selection model. Therefore, I exclude the volatility factor and estimate the model using the nine remaining factors for all of the empirical work in this paper.

Table 3 provides the time-series means of the monthly premia. Because the model is estimated monthly while the exposures draw upon less frequently updated data, the standard errors incorporated in the t-statistics are calculated using the method of Newey and West (1987) and allow for 11 lags of overlap. The estimated premia are mostly consistent with the anomalies literature that motivates the factors. There is a clear value premium attached to the earnings yield, dividend yield, book-to-market, and profitability. There is also an illiquidity premium attached to lower dollar volume. A momentum strategy generates a positive average return during this sample, despite the large negative returns it produced in the early portion of 2009. The premium normally attached to small firms is actually negative but only marginally significant. Market beta and the debt ratio have signs opposite to what is expected, although neither is significant As suggested by Table 3, there is substantial time variation in the estimated premia, and the decade from 2000–2010 witnessed especially high volatility for many of these standard return anomalies.

The goal of the multifactor model used in this research is to capture the investment signals on which many investors may have traded. As an informal but useful check on the reasonableness of the model and its ability to achieve this goal, I examine the factor premia during July and August 2007. During this time, the quantitative equity approach experienced severe upheaval as stock selection factors generated strongly perverse performance, causing many portfolio managers employing these models to experience significant short-term losses. Rothman (2007), Khandani and Lo (2007), Asness (2008), Khandani and Lo (2011), and Cahan and Luo (2013) discuss this episode in detail. One pertinent aspect of this episode is that while there was nothing extraordinary

[4] Some studies support the notion that idiosyncratic risk earns a positive risk premium on average, notably Goyal and Santa-Clara (2003), whereas other articles such as Guo and Savickas (2008), Pollet and Wilson (2010), and Chen and Petkova (2012) document a negative relation between idiosyncratic risk and returns.

about the performance of the aggregate market during this time, multifactor models (in particular, traditional valuation factors) experienced a systematic disruption of historic magnitude that was largely reversed by the end of August. To be considered an accurate proxy for the great variety of multifactor models in use, the model employed in this paper should at a minimum exhibit this pattern during the summer of 2007.

Figure 1 depicts the performance of two valuation factors included in the multifactor model. To create this plot, the multifactor model is estimated using daily returns as the dependent variable and the resulting daily estimates of the factor premia are compounded over an interval beginning on July 23, 2007, and ending on August 31, 2007. For reference, the plot also includes the cumulative performance of the Fama-French HML portfolio as another measure of the return to a value strategy. The results illustrate the large negative factor returns which betting on value stocks generated during this period. The losses associated with HML are larger than the other two factors, which is not surprising as its method of construction controls only for firm size as a competing factor, whereas the factor premia are estimated within a multivariate regression framework. It is worth noting that the two negative premia are additive to the extent that book-to-market and earnings-to-price exposures are correlated across stocks. Aside from the larger loss, HML shows only a small recovery by the end of August, whereas the other two factors recover much of the lost ground. The behavior of earnings-to-price and book-to-market depicted in Figure 1 is consistent with the experiences that many investment managers reported at the time, and provides confirmation that the specification of the multifactor model used in this paper overlaps with the stock selection factors that investment managers employ.[5]

2.2 Out-of-Sample Performance of the Multifactor Model

At the end of each month in the sample period, the multifactor model is used to generate expected returns on all stocks in the sample. Expected returns are calculated as the product of the most recent observations of standardized factor exposures and the historical average factor premia. In particular, at the end of month t, a twelve-month moving average of the premia estimated using

[5] Rothman (2007) was published on August 9, 2007, and provides an immediate response to portfolio managers' worries about the magnitude of the perverse factor returns during the preceding days. The main purpose of the report is to publicly recognize that the disruption was affecting the entire quantitative equity industry and was not a failure of any particular model.

(1) for $\{t-12, t-11, \ldots, t-1\}$ is used as the historical average. Once expected returns have been calculated, the sample of stocks is ranked and divided into quintiles in which quintile 5 represents the stocks with the highest expected returns and quintile 1 contains the stocks with the lowest expected returns.

Table 4 provides average returns to portfolios based on the expected return signals, and thus a measure of how the multifactor model performs out-of-sample. Average returns are strongly related to the expected return generated by the model. The top row of the table shows that an equal-weighted long-short strategy of holding the highest expected return quintile and shorting the lowest expected return quintile produced an average monthly return of 0.97%. Looking more generally at the entire cross-section, the time-series mean of the Spearman rank correlation between expected returns and subsequent realized returns (the rank information coefficient) is 0.056. The partial signals in the following rows correspond to independent strategies that select stocks only on the basis of exposures to the indicated factor. Small stocks, high beta stocks, momentum winners, value stocks, and illiquid stocks all offer higher average returns. Despite the Quant Crisis at the beginning of August 2007, using book-to-market as a long-short signal yields a highly significant monthly return of 1.12%, and book-to-market is the most powerful factor during this period (owing in part to its recovery in the second half of August 2007, which makes the episode seem small when using monthly data). On the whole, the results in Table 4 are consistent with directional predictions of the factors, the magnitude of the gross returns documented in the anomalies literature, and the level of predictability one might expect from a multifactor model of this type.

3 Main Results

This section presents the main empirical findings supporting the notion that many investors use similar investment signals, and that the unintended coordination of their trading both exhausts the liquidities of the stocks they trade and alters their liquidity risk. To make this argument, Section 3.1 first describes how the multifactor stock selection model developed in the previous section is used to generate trade signals. Patterns in the level of share turnover and stock volatility surrounding the point in time in which the signal is generated are studied in Sections 3.2 and 3.3. These patterns are consistent with investors acting upon the signals and trading in a directional manner to cause an order imbalance. An order imbalance would tend to strain available liquidity, and consequently Section 3.4 examines changes in liquidity and documents that on average liquidity decreases substantially following a signal. Finally, recognizing that such a model is typically used to rank and trade many stocks concurrently, Section 3.5 examines how stocks traded by the model load on liquidity risk factors before and after the signal. This section presents evidence that after a signal is generated for a stock, its liquidity covaries less with an aggregate liquidity factor and more with the liquidities of other stocks traded by the model. This is consistent with the stocks favored by the model being traded as a basket, and demonstrates that overlapping investment models are one source of commonality in liquidity.

3.1 Trade Signals

The multifactor model studied in Section 2 is a parsimonious example of a stock selection model intended to be as general and as uncontroversial as possible. It relies on standard factors that have been studied extensively in the anomalies literature and these factors have been processed in the standard way to produce expected return forecasts. Because many investors incorporate proprietary research into their investment process, the expected return rankings generated by this simple model should not be used as reliable trade indicators. Furthermore, knowing the relative ranking of a stock may not provide much information about the timing of a trade, as would be the case for a stock that ranked highly throughout a long period. Instead of trying to capture trade signals directly through the level of expected return, I focus on significant changes in expected return. If two models rely on similar variables, but the precise manner in which the factors are computed and the weights applied to them are different enough, they may generate significantly different rankings for a sample of stocks. However, when a multifactor model abruptly changes the expected return from low to high or vice versa, it is likely due to simultaneous changes in the exposures to multiple factors. If two models use similar factors and at least agree on the signs of the weights, these extreme exposure changes will cause both models to revise their expected return estimates at the same time and in the same direction. If sufficiently large, the change in the expected return is a trade signal produced by both models at the same time, leading to correlated trade.

To capture large changes in expected returns, I examine transitions between quintile portfolios based on expected returns. In particular, at the end of month t, expected returns for month $t+1$ denoted by $E_t[r_{it+1}]$ are estimated for all stocks using the multifactor model, and stocks are sorted into quintiles on the basis of expected return. Let $Q(E_t[r_{it+1}])$ be the quintile ranking assigned to a stock's expected return at time t so that quintile 5 (1) contains the highest (lowest) expected return stocks. Trade signals are identified by comparing $Q(E_t[r_{it+1}])$ to $Q(E_{t-1}[r_{it}])$ for each stock in the sample using the following definitions:

Signal Type	$Q(E_{t-1}[r_{it}])$	$\rightarrow$	$Q(E_t[r_{it+1}])$
Strong Buy	1		5
Weak Buy	1		4
Weak Buy	2		5
Strong Sell	5		1
Weak Sell	4		1
Weak Sell	5		2

There are many considerations relevant to a trade decision in addition to the expected return, not least among them risk and liquidity constraints. Nonetheless, it is reasonable to expect that a stock with a large upward revision to its expected return would be the subject of increased buy activity. If this manner of classifying trade signals is too noisy, then one would not expect to document the associated changes in trade activity, volatility, and liquidity that I present in the following sections.

Table 5 provides a summary of the transition frequencies across expected return quintiles. The main diagonal of the transition matrix contains just under 64% of the sample, and the first upper and lower diagonals contain an additional 15% each. This creates the impression of relatively low volatility in revisions to expected returns, a fact that is not surprising because many of the factors are based on financial statement data that updates either quarterly or annually. Taken together, strong buys (upper-right corner) and strong sells (lower-left corner) account for only 0.30% of a sample of 1.23 million trade signals. Weak buys and weak sells sum to just over 1% of the sample. While the frequency of the trade signals on which I focus is low, it should be noted that these signals would be expected to account for a disproportionate share of the trade activity for investors using this or a similar model.

The transition matrix is a summary of trade frequencies in the pooled sample. For more detail on the output of the stock selection model, Figure 2 plots the number of each type of signal produced at the end of each month in the sample. Overall, there are 1798 strong buy signals, 1757 strong sell signals, 6738 weak buy signals, and 6280 weak sell signals. There is clearly periodic behavior in the timing of the signals related to the release of new annual data that flows into the model through updated factor exposures. However, the pattern is not completely regular owing to the other factors that are updated more frequently. In addition to the roughly annual cycle, there is a small uptick in model activity in the rebound following the Financial Crisis, from roughly March through December 2009. Aside from these two patterns evident in Figure 2, there is little pattern in the signals over time.

3.2 Trade Signals and Trading Activity

A first step to establishing the effects of correlated trade is to examine the level of trade activity associated with the signals described above. An increase in the level of trade activity is not a necessary condition for correlated trade to affect liquidity, as the correlation causes a bias in the net direction of trade that would consume more liquidity even if the overall level of trade activity does not change. However, one would expect that if the trade signals are capturing actual investment strategies, there will be some increase in volume associated with them.

To capture changes in trading activity, I measure average daily share turnover within each of the six months centered on the date the trade signal is generated. To allow for comparison across stocks with different levels of trade activity, I normalize the monthly averages by their own time-series average. More specifically, daily turnover for stock i on day s is computed as $TO_{is} = v_{is}/q_{is}$, where v_{is} and q_{is} are volume and shares outstanding. Let $\overline{TO}_{im}$ be the mean of TO_{is} during each of the months $m = t-3, t-2, \ldots, t+3$. Normalized turnover for stock i during month m is calculated as:

$$\overline{TO}^{*}_{im} = \frac{\overline{TO}_{im}}{\frac{1}{7}\sum_{u=t-3}^{t+3} \overline{TO}_{iu}} \qquad m = t-3, t-2, \ldots, t+3 \tag{3}$$

Normalized turnover is measured separately for strong buys, weak buys, strong sells, and weak sells. In this and much of the following analysis, all other signals are pooled together into a "No Signal" category. To create a balanced panel that allows for direct comparison, any signal for which I cannot compute the seven estimates of normalized turnover is excluded from the analysis of trading activity.

Figure 3 illustrates a strong and asymmetric pattern associated with the trade signals. Prior to the signal being generated, trade activity associated with buy signals is relatively low, running at approximately 80% of the average during this six-month window for strong buys and 90% of the average for weak buys. Following the buy signals, there is a large increase in trade activity that slowly decays during the next three months. Comparing the month following the signal with the month preceding the signal, the mean increase in normalized turnover for strong (weak) buys is a highly significant 35% (18%). This pattern is consistent with a sizable population of investors acting on the buy signals. In contrast, the level of trade activity associated with the sell signals is higher preceding the signal than after it, and there is a clear downward trend in turnover throughout the entire window. For strong sells, normalized turnover falls by roughly 13% in the month following the signal. Weak signals show a more sluggish response, falling only 0.7% at first, but continuing to fall so that normalized turnover during the three months following the signal averages 95% of the pre-signal level.

The asymmetric pattern with respect to buys and sells is not surprising. Nearly all investors could act on a buy signal, but not all could act on a sell signal. If an investor already owns the shares associated with a buy signal, they could purchase more, and if no shares are owned, a new position could be initiated with a purchase. In contrast, to trade on a sell signal the investor must already hold those shares long or be willing to short them. The number of investors already holding the stock long is likely to be a small subset of the investing population, and many investors are either prohibited from shorting altogether or are limited in the short positions that they can hold due to risk considerations. Therefore, not as many investors could trade on a sell signal, and we would expect the change in trading activity associated with buys to be larger. However, there is one way in which a sell signal could affect investors that do not hold the stock long and are unwilling to short it: These investors can ignore the stock, not trading it to hold it long and also not shorting it. The result would be a reduction in trade activity associated with sell signals, precisely as seen in Figure 3. The fact that the downward trend associated with sell signals begins before the trade signal may be evidence that firms slowly release negative information to the market in a manner not captured by the model exposures, for instance statements made by the management in anticipation of poor reported performance.

3.3 Trade Signals and Return Volatility

It is possible for the patterns in trade activity documented above to have no effect on stock liquidity. This could occur if, in the case of strong buy signals, an increase in the demand to acquire shares was somehow matched by an increase in the number of shares for sale. The key ingredient for producing a change in liquidity is correlated trade, and changes in the level of trade activity do not necessary reflect changes in correlated trade. One means of detecting correlated trade is to test for the effects that it would have upon stock volatility. Stock volatility is a function of order flow, and correlated trading changes the nature of order flow and should therefore produce a detectable effect if it is significant. More precisely, positively correlated trades would lower volatility because the trades tend to push from one side of the market, causing a trending in returns that is the opposite of volatility. For instance, if all trades submitted on a stock are orders to buy, then transactions will push the price steadily upward, and deviations from this mean upward trend will be relatively small.

The link between correlated trade and stock volatility can be made more rigorous by examining microstructure models of price formation. As a particular example, consider the model of Glosten and Harris (1988), which is based on transaction level data:

$$\Delta p_{it} = \alpha_i + \lambda_i v_{it}^* + \psi_i \Delta d_{it} + \varepsilon_{it} \tag{4}$$

In this model, v_{it}^* is signed volume (buyer-initiated is positive, seller initiated is negative) associated with a transaction, d_{it} is a binary trade indicator ($+1$ for buy, -1 for sell), and Δd_{it} and Δp_{it} measure the change in trade direction and transaction price relative to the previous trade. The parameter λ_i measures nonproportionate transaction costs (price impact) and ψ_i measures proportionate costs (bid-ask spread). Given a positive correlation between signed volume and Δd_{it}, the variance of price changes is:

$$\text{var}(\Delta p_{it}) = \lambda_i^2 \text{var}(v_{it}^*) + \psi_i^2 \text{var}(\Delta d_{it}) + \lambda_i \psi_i \text{cov}(v_{it}^*, \Delta d_{it}) \tag{5}$$

In normal circumstances when order flow is random and transactions are uncorrelated, trade direction will randomly bounce between buy orders (positive volume) and sell orders (negative volume), thereby creating a relatively high variance in both v^*_{it} and Δd_{it}. On the other hand, when trades are correlated, Δd_{it} is more often zero which leads to a lower variance in Δd_{it}. Correlated trade also causes the variance of v^*_{it} to be lower because many orders will be on the same side of the market, and therefore v^*_{it} will be either predominantly negative or positive. In essence, correlated trade leads to sequential and additive price impacts of the same sign based only on λ_i (trending), while random trade causes wider oscillations due to the differential bid and ask prices captured by ψ_i, and larger variance price impact due to the changing sign of v^*_{it}.

I examine changes in both the total volatility and idiosyncratic volatility of stocks for which the model generates a trade signal. Total volatility is defined as the standard deviation of daily total returns computed within a particular calendar month. Idiosyncratic volatility is measured as the standard deviation of the residuals obtained from estimating the Fama-French three factor model using daily returns within a particular calendar month. To test for a change in volatility associated with a trade signal generated at end of month t, I compare average volatility during month $t+1$ (from the end of month t through the end of month $t+1$) with volatility during month $t-1$ (from the end of month $t-2$ through the end of month $t-1$). The days during month t are excluded because the timing of the signal is uncertain. The factor exposures computed at the end of the month could be based on revisions in fundamental data released at the beginning of the month, and if so, trading on the signal (and any related volatility effects) would begin well before the end of the month, so the volatility during month t would be a poor reference point for measuring change.

Table 6 provides average changes in volatility for different sets of trade signals. The columns labeled "$\Delta\sigma$" measure the raw changes in volatility, and the columns labeled "$\Delta\sigma/\bar{\sigma}$" express this change more meaningfully as a fraction of longer-term average volatility. In particular, let σ_{it} be either total or idiosyncratic volatility measured during month t as described above. The change in volatility following a trade signal at the end of month t is $\Delta\sigma_{it+1} = \sigma_{it+1} - \sigma_{it-1}$, and average volatility during the prior year is $\bar{\sigma}_{it} = \frac{1}{12}\sum_{s=t-12}^{t-1}\sigma_{is}$. The table provides the means of $\Delta\sigma_{it+1}$ and $\Delta\sigma_{it+1}/\bar{\sigma}_{it}$ taken across the observations in each set of trade signals.

The table conveys a clear tendency for volatility to decrease following a trade signal. On average, idiosyncratic volatility decreases by 0.18%, or 3.3% of its longer term average. The magnitude of the decrease is approximately the same for strong and weak signals, although the statistical significance associated with weak signals is higher (likely owing to a substantially larger sample). Volatility decreases more following buy signals than it does for sell signals, a finding consistent with more investors being able to act upon positive signals. These same patterns in volatility changes are evident when focusing on total volatility instead of idiosyncratic volatility. In particular, volatility falls following all signals, the magnitude of the drop is roughly the same for strong and weak signals, but the drop is larger for buys than it is for sells.

3.4 Trade Signals and the Level of Stock Liquidity

The preceding sections document changes in trading volume associated with trade signals and concurrent decreases in volatility consistent with a net bias in trade direction. In addition to affecting volatility, a bias in trade direction should affect stock liquidity due to the order imbalance. In the case of a strong buy signal, there will be excess demand to purchase shares that can only be satisfied by attracting sellers through higher prices or dealer intervention. In the extreme case, each additional buy order will push the price higher and the price impact measured during this period will be large. Stock prices reacting in this manner to order imbalances is consistent with the dealer inventory paradigm modeled by Stoll (1978), in which dealer costs increase as inventory deviates from an optimal level, and hence order imbalances affect the provision of liquidity.[6]

I examine changes to Amihud liquidity surrounding the generation of a trade signal. To do so, I employ an approach similar to measuring the change in volatility in the previous section. For each signal in the sample at the end of month t, average Amihud liquidity of the stock is computed using daily data within each of the months $t-12$, $t-11$, …, $t+1$. Letting these monthly averages be denoted as AL, the change in Amihud liquidity is calculated as $\Delta AL_{it+1} = AL_{it+1} - AL_{it-1}$, where month t is again excluded to provide a clean measure of change. A longer-term pre-signal average for Amihud liquidity is computed as $\overline{AL}_{it+1} = \frac{1}{12}\sum_{s=t-12}^{t-1} AL_{is}$. Table 7 presents the percentage change in Amihud liquidity computed as $\Delta AL/\overline{AL}$ for each group of signals. On average, Amihud liquidity rises by a very significant 18.8%, signaling a strong tendency for stock liquidity to decrease

[6]Looking at the stocks traded on the NYSE in aggregate, Chordia, Roll, and Subrahmanyam (2002) find that daily order imbalances are strongly associated with lower liquidity, regardless of the sign of the bias in trade direction.

following a trade signal. The percentage decrease in liquidity is larger for strong signals than it is for weak signals, and it is also larger for sell signals than it is for buys signals. The latter effect is consistent with literature that examines asymmetry in liquidity with respect to trade direction. Brennan, Huh, and Subrahmanyam (2013) decompose Amihud liquidity based on the sign of the return and a proxy for order imbalance and find that the relation between Amihud liquidity and stock returns is strongest when trade is biased toward sells and the return is negative. This finding supports the notion that stock liquidity is especially sensitive to sell-offs, and consistent with this finding Table 7 shows the largest decrease to liquidity is associated with strong sell signals.

Table 7 also divides the signals based on the market capitalization of the stocks as ranked at the time the signal is generated. Large stocks are in the top 1000 (roughly corresponding to the Russell 1000), small/midcap stocks are ranked from 1001–3000 (roughly the Russell 2000), and microcap stocks are ranked below 3000. There are two opposing predictions concerning the relation between trade signals, the change in liquidity, and market cap. It may be the case that trade signals affect the liquidities of larger stocks more than smaller stocks because the institutional investors that are more likely to use a multifactor approach cannot feasibly trade small stocks. This is consistent with the results in Kamara, Lou, and Sadka (2008) and their conclusion that large stock liquidities covary to a greater extent in recent times due to an increase in institutional ownership. On the other hand, large stocks are held by many passive index funds, and on average they are very actively traded and enjoy a relatively high level of liquidity. Therefore, one might expect large stocks to better absorb the correlated trade resulting from overlapping models of this type, and that the liquidities of small stocks would exhibit the stronger decline. Ultimately it is an empirical question, and Table 7 provides strong evidence in favor of the latter prediction. In particular, the drop in liquidity is largest and most significant for the smallest stocks, while the change in liquidity is close to zero and generally insigificant for the largest stocks. However, even among the largest stocks there is evidence that liquidity falls following strong sell signals.

3.5 Trade Signals and Liquidity Comovement

If individual stock liquidity responds to order imbalances associated with correlated trade, and similar models generate signals for many stocks at the same time, then it is reasonable to expect that the comovement of stock liquidity will also be affected by the generation of a trade signal. The stocks favored by the model as having strong signals will experience concurrent decreases in their liquidities, regardless of whether the signal to trade is a buy or a sell. These coordinated liquidity shocks should be reflected in the tendency for the liquidities of the stocks traded by the model to move together more than they otherwise would. Artificial comovement of returns induced by investors that trade in a correlated manner is studied in a theoretical model by Barberis and Shleifer (2003) and confirmed empirically by Barberis, Shleifer, and Wurgler (2005).

To test for a change in stock liquidity comovement following a signal, I adapt the approach that Kamara, Lou, and Sadka (2008) employ to document time-varying liquidity comovement as a function of market capitalization. In the present case, liquidity comovement is measured jointly with respect to two separate liquidity factors. The first factor is a traditional aggregate liquidity factor based on all stocks not traded by the model at a given point in time. In particular, let θ_t denote the set of stocks for which a strong buy, strong sell, weak buy, or weak sell signal has been generated at the end of month t, and let $\Theta_t = \bigcup_{s=t-2}^{t} \theta_s$ be all stocks affected by such signals during the most recent three months.[7] A clean measure of the aggregate liquidity of stocks not traded by the model on day s during month t is calculated by averaging Amihud liquidity over the set of stocks not in Θ_t:

$$L_s^A = \frac{1}{\#\{i \notin \Theta_t\}} \sum_{i \notin \Theta_t} \frac{|r_{is}|}{v_{is} p_{is}} \qquad \text{for day } s \text{ in month } t \tag{6}$$

Given the number of trade signals relative to the number of firm-months in the sample, the aggregate liquidity factor L^A is very close to what previous studies have used to measure systematic liquidity. Additionally, a second daily liquidity factor is constructed only from the stocks that have been

[7]The types of signals generated by this stock selection model are valuation measures attached to abnormal performance over a medium or long-term investment horizon, and are therefore not very sensitive to when the signal is acted upon. In contrast, some quantitative stock selection strategies are used for predicting short-term returns (such as high frequency trading algorithms), and must be acted upon quickly to be valuable. It is with this in mind that I use trade signals during the recent three months to define aggregate and model liquidity. In unreported tables, the definition of Θ_t was narrowed to include only signals generated for month t ($\Theta_t = \theta_t$) and also made broader by including an additional history of three month of signals ($\Theta_t = \bigcup_{s=t-5}^{t} \theta_s$). The results in both cases are similar to those presented in Table 8 and the particular assumption used is not material.

traded by the model in recent months:

$$L_s^M = \frac{1}{\#\{i \in \Theta_t\}} \sum_{i \in \Theta_t} \frac{|r_{is}|}{v_{is} p_{is}} \qquad \text{for day } s \text{ in month } t \tag{7}$$

If a stock is traded by the model in month t, it is excluded from the calculation of L^M for months $t-1$ and t.[8] This ensures that the covariance of stocks traded by the model with L^M are not inflated simply due to their inclusion in the factor. To estimate liquidity comovement for each stock having a signal, a time-series regression of daily percentage changes in individual stock liquidity are regressed on percentage changes in the aggregate liquidity factors:

$$\ln\left(\frac{AL_{is}}{AL_{is-1}}\right) = \alpha_i + \beta_i^A \ln\left(\frac{L_{is}^A}{L_{is-1}^A}\right) + \beta_i^M \ln\left(\frac{L_{is}^M}{L_{is-1}^M}\right) + \epsilon_{is} \tag{8}$$

A specification using percentage changes in liquidity is preferred due to the fact that average Amihud liquidity falls dramatically during the sample and is therefore non-stationary. This specification also facilitates comparison across stocks that have different levels of liquidity. For each trade signal generated at time t, equation (8) is estimated during a pre-signal period defined as $[\,t - 142 \text{ trade days},\ t - 22 \text{ trade days}\,]$ and a post-signal period defined as $[\,t + 1 \text{ trade day},\ t + 120 \text{ trade days}\,]$. Any changes in the comovement of stock liquidity associated with a trade signal can be measured by comparing the estimates of β^A and β^M from the pre-signal period with those obtained during the post-signal period.

Table 8 provides the mean estimates of the liquidity betas from (8) averaged across the various sets of trade signals. The estimates make it clear that the aggregate liquidity factor L^A plays a central role in explaining individual stock liquidity. However, caution should be exercised when comparing the magnitudes of the liquidity betas because the variance of L^M is larger than the variance of L^A, as at any given time L^M averages across a much smaller set of stocks. The final four columns of the table test the mean changes in the liquidity betas. The results present a consistent picture in which the liquidities of stocks that are favored by the model begin to covary more strongly with other stocks favored by the model, and covary less with aggregate market

[8] In other words, every stock that is traded by the model has its own version of L^M in which that stock has been purged from the calculation of equation (7).

liquidity. Pooling the set of signals, the mean change in the aggregate liquidity beta β^A is negative and the mean change in β^M is positive, and both estimates are statistically significant. When examining the categories of signals individually, the pattern of changes is preserved but in many cases the estimate is not statistically significant at usual levels. In particular, the decrease in β^A following the signal is not statistically significant for strong buys and sells, and the increase to β^M is only significant for weak buys.

The magnitudes of the changes in both liquidity betas are larger for buy signals than for sell signals. This finding fits well with the evidence on trade signals and trading activity in Section 3.2. A buy signal is an active signal in the sense that it induces many investors to submit purchase orders for shares, and correlated trade and the resulting order imbalance will exist to some extent until all interested investors have acted or the signal decays. During this period, the liquidity of the stock will on average decrease, and will do so at the same time as other stocks being traded by the model. This leads to the strong and significant increases to β^M and decreases to β^A seen for buy signals in Table 8. In contrast, sell signals are much less active than buy signals. A few current shareholders may reduce their holdings based on the signal but they are likely to be a small subset of all investors, and some small number of other investors may short the stock, but the greatest effect of a sell signal may be for potential buyers to ignore the stock. In the case of a buy signal, a large population of investors flock to the market to trade in the same direction, whereas for a sell signal a small number of investors may trickle in sell orders to an inactive market. In both cases, stock liquidity will decrease, but it appears that sell signals do not induce sufficient activity to generate a significant change in the comovement of liquidity.

4 Summary and Conclusions

A decision to trade a stock should be made based on the expected net return that can be earned, and the net return is dependent on the stock's liquidity. Limitations on liquidity and the impact of trading is a primary concern for institutional investors that manage large portfolios. Recognizing the importance of liquidity and the role it plays in investment decisions, much research has been done to further our understanding of how to measure it and what determines it, both cross-sectionally and over time. While the present study shares the goal of better understanding liquidity, it is different from previous studies because of its focus on how investment decisions affect liquidity rather than vice versa. This focus embodies the idea of factor crowding, and it is an important topic for study as technological progress makes it easier for investors to implement overlapping quantitative models.

The evidence presented in this study supports the argument that overlapping equity investment strategies affect the liquidities of the stocks that they trade. Inspired by a widely used commericial multifactor quantitative stock selection model, I measure trade signals as large and abrupt changes in the expected return rankings generated by the model. I find that the stocks associated with these trade signals exhibit the patterns in trade activity and volatility that one would expect if many investors acted on the signals. Consistent with the directional nature of a signal and the resulting order imbalance, I also find that individual stock liquidity decreases following the signal. Finally, consistent with the notion that stocks favored by the model are linked by the existence of a signal, this study demonstrates that the liquidities of stocks with trade signals covary more strongly after the signal is generated, and at the same time their exposure to aggregate liquidity risk decreases.

Factor crowding most often refers to the idea that the profits associated with using common investment signals for stock selection are diminishing as more investors employ these signals. However, a focus on profitability is incomplete. Some strategies may thrive on crowding, such as momentum, and for these is factor crowding unimportant because they remain profitable? This study is focused on risk, a dimension to factor crowding that is equally important to profitability, but for which there is little empirical evidence at the present. To an investor who uses the types of signals employed in this study, the increased comovement of stock liquidity is a risk that cannot be diversified except by trading outside the model. Measuring liquidity on the basis of the stock's historical liquidity level and historical liquidity beta understates the true risk of the portfolio because it ignores the additional risk created through crowding. The consequences of this additional risk can be severe as evidenced by recent episodes in financial markets.

The obvious way to mitigate this additional liquidity risk is to decrease the overlap in positions by creating a stock selection model less correlated to the consensus. Factor crowding has the flavor of a contagion in which a concentrated liquidity event for one investor can be transmitted to other investors in concentrated form through overlap, such as occurred during August 2007. Many institutional investors allocate substantial resources to developing unique models and investment strategies, and for these investors additional liquidity risk will be lower and their exposure to factor crowding less of a concern. However, it is difficult to completely eliminate this type of risk because many multifactor models at their core are reliant on the same expected regularities in stock returns. This risk must also be managed, and to do so requires the development of portfolio risk metrics that incorporate the notion of factor crowding. I leave this challenge to future research.

References

Acharya, Viral V., and Lasse Heje Pedersen, 2005, Asset pricing with liquidity risk, *Journal of Financial Economics* 77, 375–410.

Amihud, Yakov, 2002, Illiquidity and stock returns: Cross-section and time-series effects, *Journal of Financial Markets* 5, 31–56.

——— , and Haim Mendelson, 1986, Asset pricing and the bid-ask spread, *Journal of Financial Economics* 17, 223–249.

Asness, Clifford S., 2008, The past and future of quantitative asset management, *CFA Institute Conference Proceedings Quarterly* 25, 47–55.

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

——— , and Jeffrey Wurgler, 2005, Comovement, *Journal of Financial Economics* 75, 283–317.

Brennan, Michael J., Sahn-Wook Huh, and Avanidhar Subrahmanyam, 2013, An analysis of the Amihud liquidity premium, *Review of Asset Pricing Studies* 3, 133–176.

Brennan, Michael J., and Avanidhar Subrahmanyam, 1996, Market microstructure and asset pricing: On the compensation for illiquidity in stock returns, *Journal of Financial Economics* 41, 441–464.

Brunnermeier, Markus K., and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2238.

Cahan, Rochester, and Yin Luo, 2013, Standing out from the crowd: Measuring crowding in quantitative strategies, *Journal of Portfolio Management* 39, 14–23.

Caporin, Massimiliano, Angelo Ranaldo, and Paolo Santucci de Magistris, 2013, On the predictability of stock prices: A case for high and low prices, *Journal of Banking and Finance* 37, 5132–5146.

Chen, Zhanhui, and Ralista Petkova, 2012, Does idiosyncratic volatility proxy for risk exposure?, *Review of Financial Studies* 25, 2745–2787.

Chincarini, Ludwig B., and Daehwan Kim, 2006, *Quantitative Equity Portfolio Management: An Active Approach to Portfolio Construction and Management* (McGraw-Hill Professional: New York) first edn.

Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2000, Commonality in liquidity, *Journal of Financial Economics* 56, 3–28.

——— , 2002, Order imbalance, liquidity, and market returns, *Journal of Financial Economics* 65, 111–130.

Easley, David, and Maureen O'Hara, 1987, Price, trade size, and information in securities markets, *Journal of Financial Economics* 19, 69–90.

Edirisinghe, N.C.P., and X. Zhang, 2007, Generalized DEA model of fundamental analysis and its applicaton to portfolio optimization, *Journal of Banking and Finance* 31, 3311–3335.

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Feldman, R., S. Govindaraj, J. Livnat, and B. Segal, 2010, Management's tone change, post earnings announcement drift and accruals, *Review of Accounting Studies* 15, 915–953.

Glosten, Lawrence R., and Lawrence E. Harris, 1988, Estimating the components of the bid/ask spread, *Journal of Financial Economics* 21, 123–142.

Goyal, Amit, and Pedro Santa-Clara, 2003, Idiosyncratic risk matters!, *Journal of Finance* 58, 975–1007.

Grossman, Sanford J., and Merton H. Miller, 1988, Liquidity and market structure, *Journal of Finance* 43, 617–633.

Guo, Hui, and Robert Savickas, 2008, Average idiosyncratic volatility in g7 countries, *Review of Financial Studies* 21, 1259–1296.

Gustafson, Keith, and Patricia Halper, 2010, Are quants all fishing in the same small pond with the same tackle box?, *Journal of Investing* 19, 104–115.

Harris, Lawrence, and Eitan Gurel, 1986, Price and volume effects associated with changes in the s&p500 list: New evidence for the existence of price pressures, *Journal of Finance* 41, 815–829.

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

——— , and Duane J. Seppi, 2001, Common factors in prices, order flows, and liquidity, *Journal of Financial Economics* 59, 383–411.

Jacobs, Bruce I., and Kenneth N. Levy, 1988, Disentangling equity return regularities: New insights and investment opportunities, *Financial Analysts Journal* 44, 18–43.

Kamara, Avraham, Xiaoxia Lou, and Ronnie Sadka, 2008, The divergence of liquidity commonality in the cross-section of stocks, *Journal of Financial Economics* 89, 444–466.

Khandani, Amir E., and Andrew W. Lo, 2007, What happened to the quants in august 2007?, *Journal of Investment Management* 5, 5–54.

———, 2011, What happened to the quants in august 2007? evidence from factors and transactions data, *Journal of Financial Markets* 14, 1–46.

Lo, Andrew W., and Jiang Wang, 2000, Trading volume: Definitions, data analysis, and implications of portfolio theory, *Review of Financial Studies* 13, 257–300.

McKay Price, S., James S. Doran, David R. Peterson, and Barbara A. Bliss, 2012, Earnings conference calls and stock returns: The incremental informativeness of textual tone, *Journal of Banking and Finance* 36, 992–1011.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 817–858.

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and price discovery, *Journal of Political Economy* 111, 642–685.

Pollet, Joshua M., and Mungo Wilson, 2010, Average correlation and stock market returns, *Journal of Financial Economics* 96, 364–380.

Rothman, Matthew, 2007, Turbulent times in quant land, Lehman Brothers Research.

Shleifer, Andrei, 1986, Do demand curves for stocks slope down?, *Journal of Finance* 41, 579–590.

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

———, 1989, Inferring the components of the bid-ask spread: Theory and empirical tests, *Journal of Finance* 44, 115–134.

Tetlock, Paul C., 2007, Giving content to investor sentiment: The role of media in the stock market, *Journal of Finance* 62, 1139–1168.

Verbeek, Marno, and Yu Wang, 2013, Better than the original? The relative success of copycat funds, *Journal of Banking and Finance* 37, 3454–3471.

Werner, Antweiler, and Murray Z. Frank, 2004, Is all that talk just noise? The information content of internet stock message boards, *Journal of Banking and Finance* 59, 1259–1293.

Table 1: Factor Definitions and Data Sources

This table provides a description of each of the predictive variables used in the multifactor model. These variables are motivated by those employed by the Barra USE3 Equity Risk Model and are chosen to capture the major risk indexes of that model in a simple way. Accounting data is obtained from the quarterly and annual Compustat Industrial databases and market data is obtained from CRSP.

Barra Risk Index	Factor Construction (Descriptor)
Beta	Slope coefficient on the CRSP value-weighted index from estimating a regression using the most recent five years of monthly data, requiring at least 36 valid return observations.
Momentum	12-month momentum computed as the cumulative total return earned during the eleven-month period ending one month prior to exposure calculation.
Size	Log of market capitalization calculated using daily closing prices and shares outstanding from one trading day prior to exposure calculation.
Earnings/Price	Most recently reported annual EPS allowing a 90 calendar day reporting lag, divided by daily price one day prior to exposure calculation.
Volatility	Variance of residual terms relative to the Fama-French three-factor model estimated using one year of daily returns ending on the day prior to exposure calculation.
Growth	Gross profit margin calculated using revenue and cost of goods sold summed over the most recent four quarters, allowing 60 calendar days for a reporting lag.
Dividend Yield	The sum of dividends paid during the most recent annual period divided by the closing price on the day prior to exposure calculation.
Book-to-Price	The book-to-market ratio calculated as in Fama and French (1996).
Trading Activity	Average dollar volume computed using daily observations of closing price and shares traded during the month ending on the day prior to exposure calculation.
Leverage	Debt ratio computed as long-term debt to total assets based on quarterly statement data allowing for a 60 calendar day reporting lag.

Table 2: Summary of Liquidity and Factor Exposures

Exposures to factors detailed in Table 1 are calculated for all stocks in the merged CRSP/Compustat database at the end of each month for the period January 1995 through December 2014. Coverage is the percentage of the sample for which it is possible to compute an exposure. Observations of idiosyncratic volatility and market beta are considered valid only when the regressions from which they are obtained include at least 100 observations. Observations of Amihud liquidity and daily dollar volume are averages computed during the month ending on the day prior to exposure calculation, and must be based on at least 15 observations. The below statistics are the time-series means of the cross-sectional statistics indicated in the column headings.

Factor	coverage	equal-weighted mean	value-weighted mean	median	standard deviation
Idiosyncratic volatility (%)	96.61	4.0917	1.9246	2.9431	4.7810
Gross profit margin	93.62	0.0150	0.4302	0.3688	2.4754
12-month momentum	95.86	0.0842	0.0875	0.0195	0.3397
Earnings yield	93.92	-0.0951	0.0380	0.0360	0.4349
Dividend yield	99.22	0.0098	0.0179	0.0000	0.0192
Book-to-Market	94.35	0.7841	0.4179	0.5554	0.9272
Debt ratio	97.47	0.1614	0.1734	0.0923	0.1909
Market beta	93.14	1.1007	0.9933	0.9575	0.8003
Market capitalization ($B)	99.22	2.1277	26.7902	0.2671	6.3332
Daily dollar volume ($M)	98.83	17.6539	183.4268	1.4349	48.3143
Monthly return (%)	98.42	0.9895	0.8929	0.1413	14.6280
Amihud liquidity ($\times 10^6$)	93.41	0.2154	0.0020	0.0047	0.8473

Table 3: In-Sample Fitting of Multifactor Model

Each month, total stock returns earned during month t are regressed on the standardized factor exposures measured at the end of month $t-1$:

$$r_{it} = \alpha_t + \sum_{k=1}^{9} \lambda_t^k z_{it-1}^k + \epsilon_{it}$$

The below table provides the time-series means and t-statistics associated with the estimated monthly factor premia. The GLS method weights the cross-sectional observations by the inverse of residual variance measured relative to the Fama-French three-factor model based on daily data during the previous year. The t-statistics have been calculated using the method of Newey and West (1987), allowing for eleven periods of overlap based on the use of annual financial data. The sample period is 199501–201412.

	OLS		GLS	
Factor	$\bar{\lambda}$	t-stat	$\bar{\lambda}$	t-stat
Intercept	0.747	2.02	0.990	3.03
Market cap	0.134	0.96	-0.035	-1.16
CRSP-VW β	0.121	1.18	0.164	1.50
12-month momentum	0.273	2.81	0.260	2.42
Earnings/Price	0.311	3.96	0.269	2.84
Dividend yield	0.279	1.98	0.217	1.29
Dollar volume	-0.089	-1.25	-0.047	-1.96
Book-to-Market	0.279	5.53	0.158	3.00
Gross profit margin	0.272	4.39	0.395	5.02
Debt ratio	-0.064	-1.43	-0.099	-2.87
mean R^2	0.059			
median R^2	0.042			
Number of months	240			

Table 4: Multifactor Model Out-of-Sample Performance

Each month, a multifactor model of stock returns is used to generate expected returns for the following month:

$$E_t[r_{it+1}] = \sum_{k=1}^{9} z_{it}^k \bar{\lambda}_t^k \qquad \text{where} \qquad \bar{\lambda}_t^k = \sum_{s=t-12}^{t-1} \lambda_s^k \tag{9}$$

where z_{it}^k is stock i's exposure to factor k at the end of month t and λ_s^k is the estimated premium associated with factor k during month s. The Rank IC is the Spearman rank correlation between the expected return generated by the model and the subsequent monthly return. The full model uses all factors whereas partial signals are based only on the exposure to the indicated factor. The sample period is 199501–201412.

		Quintile Returns					Q1−Q5	
$E[r]$ Signal	Rank IC	Q1	Q2	Q3	Q4	Q5	return	t-stat
Full Model	0.056	0.604	0.731	1.007	1.239	1.573	0.969	2.306
Partial Signals								
Market cap	0.052	1.844	1.137	1.037	1.113	1.003	-0.840	-2.152
Market beta	0.022	1.089	1.302	1.440	1.405	1.468	0.379	1.799
Momentum	0.035	1.093	1.229	1.181	1.182	1.529	0.436	1.610
Earnings yield	0.062	0.903	1.041	0.995	1.106	1.290	0.388	1.764
Dividend yield	0.031	0.969	1.043	1.076	1.128	1.256	0.287	0.872
Dollar volume	0.037	1.571	1.198	1.179	1.191	0.981	-0.590	-1.937
Book-to-market	0.029	0.840	0.923	1.206	1.344	1.961	1.122	3.305
Gross profit margin	0.032	0.924	1.146	1.300	1.288	1.572	0.648	3.611
Debt ratio	0.012	1.455	1.171	1.174	1.250	1.059	-0.397	-1.795

Table 5: Trade Signals

At the end of each month, all stocks in the sample are ranked into quintiles on the basis of the expected returns generated by the multifactor model. $Q(E_t[r_{t+1}]) = 5$ (1) is the quintile of highest (lowest) expected return stocks when forecast at time t. The quintile ranking for each stock is compared to its ranking the previous month to determine trade signals. The below matrix provides transition frequencies for all possible trade signals using the pooled sample of signals. A stock is considered a strong buy at time t if it is in quintile 5 at the end of month t but was in quintile 1 at the end of month $t-1$. Similarly, a transition from quintile 5 at $t-1$ to quintile 1 at t would be considered a strong sell. A weak buy can be either a transition from quintile 1 to quintile 4, or a transition quintile 2 to quintile 5. A weak sell is defined symmetrically as either a transition from quintile 5 to quintile 2, or from quintile 4 to quintile 1.

Transition Frequencies %, (N=1,231,906)

		$Q(E_t[r_{t+1}])$				
		1	2	3	4	5
$Q(E_{t-1}[r_t])$	1	15.93	3.09	0.59	0.24	0.15
	2	3.20	11.52	4.00	0.96	0.31
	3	0.58	4.08	10.33	4.20	0.81
	4	0.22	0.93	4.18	10.96	3.71
	5	0.15	0.29	0.77	3.66	15.13

Table 6: Trade Signals and Changes in Volatility

Volatility is measured before and after a trade signal is generated by the multifactor model. Total volatility is defined as the standard deviation of total returns during a particular calendar month. Idiosyncratic volatility is the standard deviation of residuals associated with estimating the Fama-French three factor model using daily returns within a given month. Let σ_{it} be the volatility for stock i computed during month t The columns labeled $\Delta\sigma$ measure the raw change in volatility as:

$$\Delta\sigma_{it} = \sigma_{it+1} - \sigma_{it-1}$$

Returns during month t are excluded due to the contamination that may result from uncertain timing of the trade signal. Average volatility over the previous year is calculated as $\bar{\sigma}_{it} = \frac{1}{12}\sum_{s=t-12}^{t-1}\sigma_{is}$, and the set of columns labeled $\Delta\sigma/\bar{\sigma}$ rescale the raw change in volatility to be a percentage of this longer-term average.

	Idiosyncratic Volatility (%)				Total Volatility (%)			
	$\Delta\sigma$	t-stat	$\Delta\sigma/\bar{\sigma}$	t-stat	$\Delta\sigma$	t-stat	$\Delta\sigma/\bar{\sigma}$	t-stat
All Signals	-0.176	-6.84	-3.33	-7.02	-0.200	-8.10	-4.10	-8.45
All Strong Signals	-0.196	-2.80	-3.50	-3.27	-0.253	-3.79	-4.36	-3.92
All Weak Signals	-0.167	-6.17	-3.28	-6.20	-0.187	-7.24	-4.02	-7.52
Strong Buys	-0.410	-3.86	-6.26	-3.99	-0.544	-5.31	-8.69	-5.32
Strong Sells	-0.020	-0.22	-0.87	-0.57	0.011	0.13	-0.34	-0.22
Weak Buys	-0.188	-4.83	-2.61	-3.73	-0.280	-7.31	-4.74	-6.53
Weak Sells	-0.143	-3.81	-3.98	-5.07	-0.099	-2.82	-3.37	-4.18

Table 7: Trade Signals and Changes in Liquidity Level

For a trade signal generated at the end of month t, the average daily value of Amihud (2002) liquidity is calculated within months $t-12, t-11, \ldots, t+1$. Letting average daily liquidity for month t be denoted by AL_{it}, the raw change in liquidity associated with the signal is measured as $\Delta AL_{it+1} = AL_{it+1} - AL_{it-1}$, where observations during month t are excluded due to uncertainty in the timing of the signal. A longer-run average for Amihud liquidity is calculated as $\overline{AL}_{it} = \frac{1}{12}\sum_{s=t-12}^{t-1} AL_{is}$. The below table presents the means of $\Delta AL_{it+1}/\overline{AL}_{it}$ within the sets of trade signals.

	Full Sample		Large Caps		Small/Midcap		Microcaps	
	mean	t-stat	mean	t-stat	mean	t-stat	mean	t-stat
All Signals	0.188	15.86	-0.008	-1.15	0.062	6.96	0.389	13.25
All Strong Signals	0.259	7.55	0.024	1.32	0.055	2.29	0.488	6.93
All Weak Signals	0.173	14.35	-0.012	-1.70	0.069	7.10	0.349	11.31
Strong Buys	0.166	3.63	0.001	0.07	0.049	2.51	0.405	3.77
Strong Sells	0.358	6.49	0.060	1.79	0.095	2.99	0.572	5.59
Weak Buys	0.143	8.72	-0.053	-1.71	0.014	1.09	0.400	8.61
Weak Sells	0.196	10.85	0.037	3.43	0.125	8.71	0.320	7.27

Table 8: Trade Signals and Changes in Liquidity Comovement

Two daily liquidity factors are formed each month. The first factor, L^A, is an aggregate liquidity factor that averages Amihud liquidity across all stocks that have not been chosen for trade by the model during the current or previous two months. The second factor, L^M, specifically measures the liquidities of stocks for which the model has generated some trade signal (strong buy, strong sell, weak buy, or weak sell). Percentage changes in individual stock liquidity are regressed on a constant and both liquidity factors:

$$\ln\left(\frac{AL_{is}}{AL_{is-1}}\right) = \alpha_i + \beta_i^A \ln\left(\frac{L_{is}^A}{L_{is-1}^A}\right) + \beta_i^M \ln\left(\frac{L_{is}^M}{L_{is-1}^M}\right) + \epsilon_{is}$$

For a signal generated at time t, estimates during the pre-signal period are obtained by using daily data in the interval $[\,t-142$ trade days, $t-22$ trade days$\,]$ and post-signal estimates rely on data during $[\,t+1$ trade day, $t+120$ trade days$\,]$. The statistics in the first six columns are means taken across the set of indicated signals.

	Pre-Signal			Post-Signal			Change: Pre-Signal − Post-Signal			
Signal Type	β^A	β^M	R^2	β^A	β^M	R^2	$\Delta\beta^A$	$\Delta\beta^M$	$t(\Delta\beta^A)$	$t(\Delta\beta^M)$
All Signals	0.295		0.021	0.268		0.021	-0.027		-2.319	
	0.292	0.014	0.038	0.265	0.023	0.040	-0.027	0.009	-2.246	3.607
All Strong Signals	0.273		0.020	0.250		0.022	-0.022		-0.859	
	0.268	0.009	0.037	0.248	0.020	0.041	-0.020	0.011	-0.769	2.065
All Weak Signals	0.301		0.021	0.273		0.021	-0.028		-2.168	
	0.298	0.015	0.038	0.270	0.024	0.039	-0.028	0.008	-2.131	2.992
Strong Buys	0.254		0.021	0.214		0.019	-0.040		-1.069	
	0.250	0.010	0.039	0.213	0.022	0.038	-0.037	0.012	-0.979	1.821
Weak Buys	0.292		0.022	0.249		0.020	-0.043		-2.317	
	0.286	0.013	0.040	0.246	0.024	0.039	-0.040	0.011	-2.079	2.709
Strong Sells	0.290		0.019	0.284		0.024	-0.006		-0.156	
	0.285	0.008	0.035	0.281	0.019	0.044	-0.004	0.011	-0.121	1.708
Weak Sells	0.310		0.020	0.298		0.021	-0.012		-0.688	
	0.311	0.018	0.036	0.295	0.023	0.039	-0.016	0.005	-0.887	1.434

Figure 1: Valuation Factors During the August 2007 Quant Crisis

The below plot illustrates the performance of two valuation factors during late-July and August of 2007. The plots associated with earnings-to-price and book-to-market are created by estimating the multifactor model (Equation (1)) using daily total returns as the dependent variable, and then compounding the regression coefficients over the interval of time depicted. For reference, the plot include the cumulative return to the Fama-French HML factor which is compounded in a similar manner.

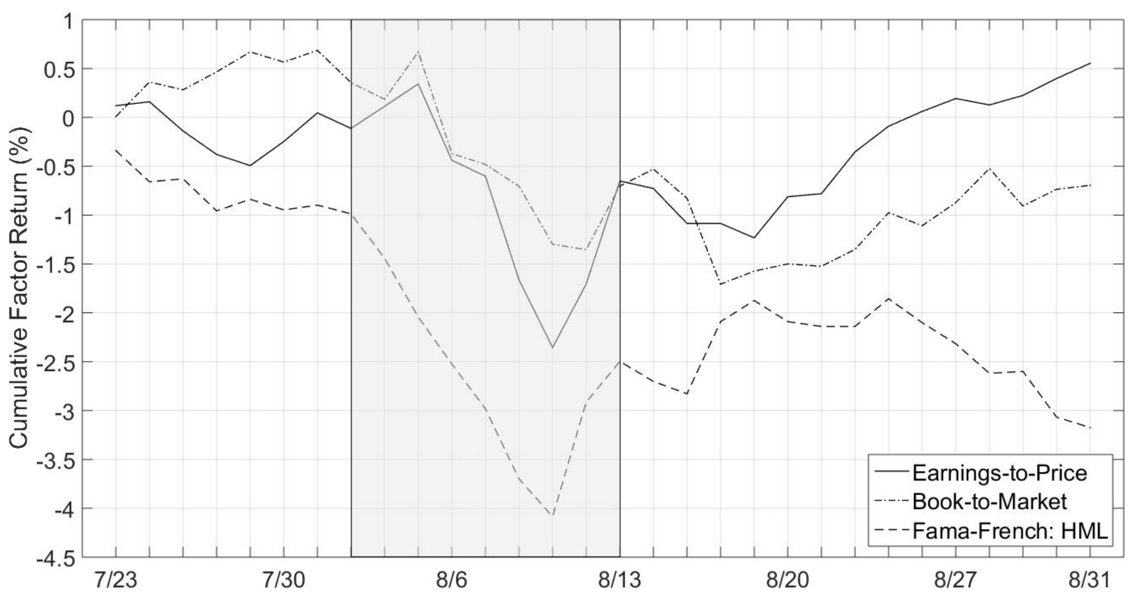


Figure 2: Distribution of Trade Signals During the Sample Period

Trade signals are generated for all stocks in the sample at the end of each month as described in Table 5. The below figure presents the number of signal of each type generated for each month in the sample period, 199501 – 201412.

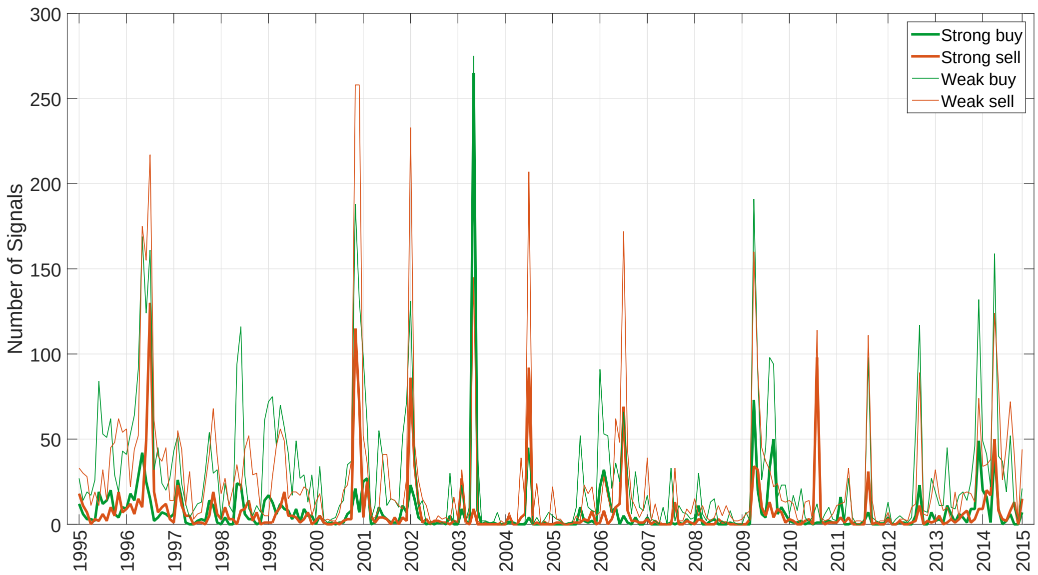
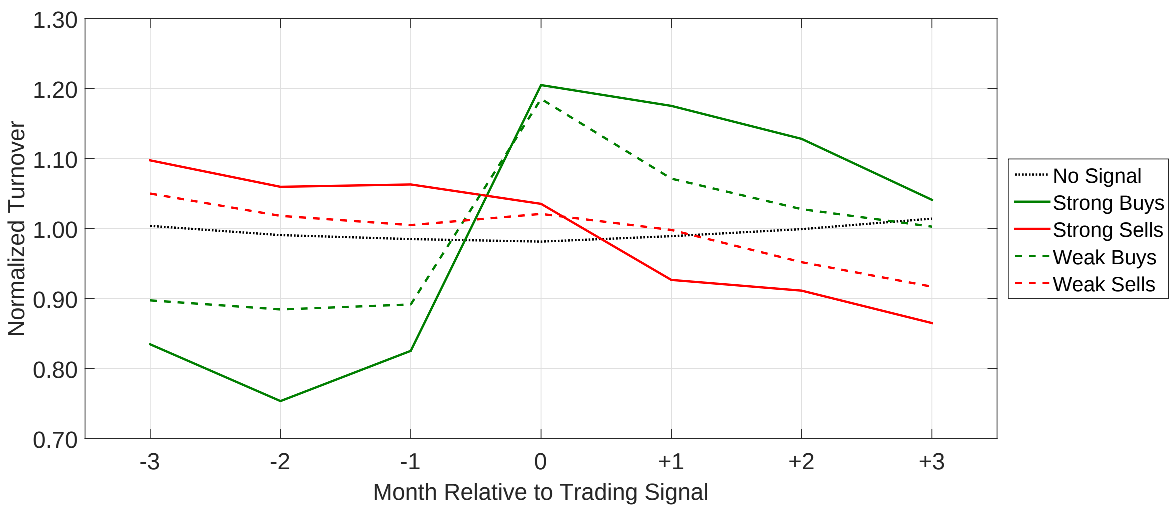


Figure 3: Trade Signals and Volume

Trade signals are identified as transitions between expected return quintiles as described in Table 5. For a signal generated at the end of month t, I measure average daily turnover (volume/shares outstanding) within months $t-3, t-2, \ldots, t+3$. Let the average daily turnover for stock i during month t be denoted as $\overline{TO}_{it}$. To facilitate comparison of stocks with different levels of turnover, the monthly averages are normalized by their time-series average:

$$\overline{TO}^*_{it} = \frac{\overline{TO}_{it}}{\frac{1}{7}\sum_{s=t-3}^{t+3} \overline{TO}_{is}}$$

Normalized turnover is calculated for strong and weak buys and sells, and all other signals are pooled together into a single group. The plot depicts the average at each point in time across signals in the indicated group. The column labeled $\Delta\overline{TO}^*$ measures the change in normalized volume as $\overline{TO}^*_{t+1} - \overline{TO}^*_{t-1}$ for each group.



	$t-3$	$t-2$	$t-1$	t	$t+1$	$t+2$	$t+3$	$\Delta\overline{TO}^*$	t-stat
No Signal	1.004	0.990	0.985	0.981	0.989	0.999	1.014	0.004	1.079
Strong Buys	0.834	0.753	0.825	1.205	1.175	1.128	1.041	0.350	3.947
Strong Sells	1.097	1.059	1.063	1.035	0.926	0.911	0.865	-0.136	-2.161
Weak Buys	0.897	0.884	0.891	1.185	1.071	1.027	1.003	0.180	6.467
Weak Sells	1.050	1.018	1.005	1.021	0.998	0.952	0.917	-0.007	-0.661