2012

# Trading Volume and Overconfidence with Differential Information and Heterogeneous Investors

Kenneth Yung
*Old Dominion University*, kyung@odu.edu

Qian Sun

Hamid Rahman



Original Publication Citation

Yung, K., Sun, Q., & Rahman, H. (2012). Trading volume and overconfidence with differential information and heterogenous investors. *Journal of Behavioral Studies in Business, 5*, 77-106.

# Trading volume and overconfidence with differential information and heterogeneous investors

Kenneth Yung
Old Dominion University

Qian Sun
Kutztown University

Hamid Rahman
Alliant International University

**ABSTRACT**

This paper adds to the overconfidence literature by specifically considering the differential nature of information and its use by different classes of investors. The literature suggests that overconfidence is a major determinant of stock trading volume. We postulate that private investors are more prone to overconfidence bias as compared to institutional investors. This implies that turnover in firms with low institutional ownership will be driven more by private information while turnover in firms with high institutional ownership will be driven more by public information. This is the essence of the two hypotheses we explore. We find strong evidence in support of the first proposition but only mixed evidence in support of the second proposition. However, the second proposition is found to be very significant in the most recent period if certain low value or low liquidity stocks are excluded from the data.

Keywords: Overconfidence, Behavioral finance, Over (under) reaction, Trading volume

**INTRODUCTION**

Stock trading volume in financial markets has been extensively studied in the literature (for example, Glaser and Weber (2007), Statman et al (2006), Chae (2005), and Covrig and Ng (2004)). This attention to trading volume is not without reason. Several studies indicate that the trading volume in financial markets far exceeds what would normally be expected from rational traders. Glaser et al (2004) calculate the trading volume as a percentage of market capitalization in 2002 to be 100% in USA, 215% in UK, 180% in Germany, 115% in France and 70% in Japan. They contend that rational motives for trade are not sufficient to explain the high trading volume. The high trading volume is even more surprising in view of the finding that those who trade the most lose the most (Odean (1999). DeBondt and Thaler (1995) observe that the high trading volume in financial markets "is perhaps the single most embarrassing fact to the standard finance paradigm."

The traditional neoclassical models of the standard finance paradigm assume investor rationality and homogeneity. These models have not been very successful in explaining many observed financial market phenomena including the high trading volume of stocks. In an attempt to explain these anomalies, there has been a gradual but perceptible shift in the finance literature to a behaviorally based paradigm in which investors are imperfectly rational and prone to systematic biases. One such judgment bias is overconfidence. Extant cognitive psychology literature establishes that overconfidence is a pervasive trait among people. Overconfident investors overrate their ability to evaluate securities as a result of overestimating the precision of their private information signals. Odean (1999) argues that overconfidence is the key determinant of trading volume. Several other researchers support this finding, for example Benos (1998), Wang (1998), Gervais and Odean (2001) and Statman et al (2006).



Psychologists also find that people systematically underweight some types of information and overweight others, and the effects of overconfidence depends on how information is distributed in a market and who is overconfident (Odean (1998)). It follows that in order to fully gauge the impact of overconfidence on stock trading volume, it is not sufficient to merely establish an aggregative relationship but also to fine tune the research format to account for the informativeness of the news and the type of investor. As the subsequent literature review shows, several researchers have considered these additional dimensions of overconfidence in isolation but never together to empirically test the relationship between overconfidence and the volume of stock trading. In this paper, we consider in the same model the differential impact of information - private and public - on investors - individual and institutional.

Trading volume arises from changes in investor beliefs associated with new information. The new information may be private or public. Daniel et al (1998) postulate that investors overreact to private information and under react to public information. Informed investors receive noisy signals about the true value of a security. If the signal is private, they react to the signal with overconfidence by overestimating its precision. If the signal is public, then investors are not overconfident and correctly estimate its precision. Chuang and Lee (2006) also find that if investors are overconfident, they overreact to private information and under react to public information. He and Wang (1995) develop a theoretical model in which they show the differential impact of public and private information on trading volume pattern. They posit that private information mostly influences the trading behavior of institutional investors while public information influences the trading behavior of both institutional and individual investors. Thus, the differential impact of private and public information on volume is well recognized in the literature and must be modeled for a comprehensive study of the effect of overconfidence on trading volume.

The differences in the trading behavior of institutional and individual investors are well documented though the findings are by no means unanimous. Barber and Odean (2008) posit that the buying and selling behavior of institutional and individual investors is different. Shefrin (2005) surveys the extant literature on heterogeneity in the judgments of individual investors and professional investors, and concludes that the two react in opposite ways to past market movements and by more than is justified. By and large individual investors forecast future returns by engaging in trend following and predicting continuation. Professional investors, on the other hand, believe they face mean reverting random processes and are excessive in predicting reversals. Covrig and Ng (2004) find a stronger relation between volume and lagged absolute return in stocks with greater institutional ownership. Cho and Jo (2006) assume that individual investors are more overconfident relative to institutions i.e. these investors are more susceptible to the psychological biases when they are processing information than institutional traders. Glaser et al (2004), however, come to the opposite conclusion. They find that judgments of professionals (traders who work in the trading room of a large bank and investment bankers) are biased, and their degree of overconfidence is higher than the respective scores of a student control group. In most tasks, this difference is significant. In our model, we let the data decide which of the two groups is more overconfident.

Tests for the empirical validation of the overconfidence theory have followed one of two tracts. One approach is to test the validity of the assumptions on which the theory is based (for example, Kirchler and Maciejovsky (2002), Hilton (2001) and Graham and Harvey (2002)) and the other is to test its predictions (for example, Statman et al. (2003), Odean (1999), Daniel et al (1998)). The most important prediction of the theory is that trading volume increases with an increasing degree of overconfidence. Odean (1998) calls this the most robust effect of overconfidence. Statman et al. (2003) test the trading volume predictions of formal overconfidence models in the U.S. stock market. Their hypothesis is that high returns will be followed by high trading volume because the investment success of investors will increase their degree of confidence. They find that share turnover is positively related to lagged returns for many months. They interpret their results as evidence of overconfidence. We follow Statman et al. to test the relationship between stock turnover and overconfidence for institutional and individual investors while allowing for the differential reaction to private and public information.

Stock prices reflect both public and private information but the relative proportion of each may differ between stocks for reasons such as the dissimilar cost of producing private information. Although it is difficult to disaggregate the two kinds of information and measure each directly, the literature suggests two indirect measures for quantifying private and public information. The first uses price nonsynchronicity as a measure of private information. It was proposed by Roll (1988). The correlation of stock return with the market and industry return is a measure of public information while the firm specific return or idiosyncratic risk is a measure of private information. Roll (1988) showed that price nonsynchronicity has very little correlation with public news and seems to capture private information. Price nonsynchronicity as a measure of private information has been used in several studies, for example Chen et al. (2007), Durnev et al. (2004), and Morck et al. (2000). The second measure of private information is Probability of Informed Trading (PIN). This measure was proposed by Easley et al. (1996). It is based on a structural market microstructure model and captures the probability of informed trading in a stock. In this study, we use price nonsynchronicity to measure private information.

The market microstructure literature recognizes the diversity of motivation, strategies and tactics of traders and accordingly models this heterogeneity by various classifications such as, ‘informed’ and ‘uninformed’ traders, ‘newswatchers’ and ‘momentum’ traders, ‘fundamentalist’ and ‘technical’ traders, ‘rational’ and ‘noise’ traders. These classifications are not mutually exclusive and often overlap to a great extent. In order to be consistent with the previous research on overconfidence, we model the trader heterogeneity in terms of ‘institutional’ and ‘individual’ traders. We start with all the New York Stock Exchange (NYSE) firms in the CRSP database and use two methods to sort the firms in accordance with their degree of institutional ownership. The first is to sort the firms into deciles; the second is to sort them in quintiles. Table 2 reports the deciles results and Table 3 the quintile results. The portfolios are rebalanced each year. We select the top and bottom decile to represent stocks with high and low institutional ownership respectively. Our model is then run separately on the high and low institutional portfolio.

The major contribution of this paper is threefold. First, it directly tests the Daniel et al. (1998) proposition that investors overreact to private information and under react to public information. Existing studies on investor overconfidence typically use stock returns as a measure of aggregate information flow without differentiating between private and public information as, for example, Statman et al(2006) and Corvig and Ng (2004). Chuang and Lee (2006) do test for the differential impact of private and public information but their model is driven by a very restrictive assumption under which public information shocks only trading volume and private information shocks only returns. Such an assumption is debatable and makes their conclusions fuzzy. This study is the first to use a direct measure of private information to test the implications of private information on overconfidence trading. Second, this paper is a far more detailed empirical investigation of the relationship between trading volume and overconfidence than has hitherto been conducted in the literature. Although the overconfidence literature has recognized the differential impact of private and public information on traders, and the differences in the behavior of institutional and individual investors, the two traits have not been investigated together. Third, this paper provides evidence regarding the extent of overconfidence in institutional and individual traders. It is generally assumed in the literature that the overconfidence trait is found dominantly, if not exclusively, in the individual investors. However, Glaser et al (2004), find that judgments of professionals traders are biased, and their degree of overconfidence are higher than the respective scores of a student control group. This research provides evidence to resolve this conundrum.

The rest of the paper is organized as follows: In section 2, we develop the hypotheses; Section 3 explains the methodology and formulates the models; Section 4 describes the data while Section 5 analyzes the results. Section 6 concludes the paper.

## HYPOTHESES

Barber and Odean (2000) find that individual investors trade more than can be rationally justified. Since this excessive trading does not lend itself to a rational explanation, behavioral models have been adduced to explain this observed market phenomenon. Statman et al. (2006) propose investor overconfidence as a major driver of over trading. Daniel et al. (1998) also model overconfidence and posit that investors overreact to private information and underreact to public information. Since excessive trading is particularly associated with individual investors, we hypothesize that stocks with low institutional ownership will be more prone to overconfidence trading. This provides the rationale for our first hypothesis:

H1: Private information is a stronger driver of stock turnover in firms with low institutional ownership as compared to firms with high institutional ownership.

Our contention is that private information comprises both good information and noise. However, the noise traders are unable to differentiate between the two and trade "on noise as if it were information" (Black (1986)). Thus, even if institutional traders are more informed and may have more good private information, we posit that their total tradable private information set is smaller because of the large preponderance of noise in the set of the individual investors who are given more to behavioral biases and fads. Thus, Dow and Gorton (2006) opine that "A large literature argues that individual investor trading is subject to a myriad of psychological biases, and that such individuals may use various heuristics, 'popular models,' as the basis for their investment decisions." A second reason in support of the hypothesis is that "the information traders can never be sure that they are trading on information rather than noise. What if the information they have has already been reflected in prices?" (Black (1986)). Black further argues that information only provides an edge and the possession of good information is not a guarantee for a profitable trade. Taking a large position means taking on more risk. So if arbitrage is costly, there is a limit to the position that a trader will take. Thus, informed institutional owners may not be able to trade very intensively on their good information.

The primary cause for stock trade is a change in the information set of investors. The information set consists of both public and private information. If the trades of individual investors are driven relatively more by private information, we surmise that the trades of institutional investors would be driven more by public information. This provides the rationale for our second hypothesis:

H2: Public information is a stronger driver of stock turnover in firms with high institutional ownership as compared to firms with low institutional ownership.

H2 is not actually hypothesized in the overconfidence literature but appears to be a natural corollary of H1 that requires empirical validation. Since noise trading – typically an individual investor phenomenon – will be strongest in firms with low institutional ownership, the relative mix of public information trading to private information trading will be in favor of the latter. The relative mix will become more favorable for public information trading as the institutional ownership increases and the influence of noise traders wanes.

## METHODOLOGY

We modify Corvig and Ng's (2004) model and apply it using the Seemingly Unrelated Regression (SUR) technique. The model postulates that stock turnover is a function of information flow and can be represented by the function:

$V_{t+1} = a + b*V_t + c*V_t*F_t$

Where $V_t$ is detrended log Turnover and $F_t$ measures information flow in time period t. Since firms typically have both institutional and private stockholders, we sort the firms in our sample into subgroups based on institutional ownership. We use five group (quintile) and ten group (decile) sorts and test the hypothesis whether private information is a stronger driver of stock turnover in firms with low institutional ownership as compared to firms with high institutional ownership by using the two groups at each end i.e. the group with the highest (hi) and the group with the lowest (lo) institutional ownership. Our two core models for testing this hypothesis take the following form:

Model 1: $V_{hi,t+1} = a_{hi} + b_{hi}*V_{hi,t} + c_{hi}*V_{hi,t}*PI1_{hi,t}$
: $V_{lo,t+1} = a_{lo} + b_{(lo)}*V_{lo,t} + c_{lo}*V_{lo,t}*PI1_{lo,t}$

Model 2: $V_{hi,t+1} = a_{hi} + b_{hi}*V_{hi,t} + c_{hi}*V_{hi,t}*PI2_{hi,t}$
: $V_{lo,t+1} = a_{lo} + b_{(lo)}*V_{lo,t} + c_{lo}*V_{lo,t}*PI2_{lo,t}$

$V_{hi(lo),t}$ is detrended logTurnover in period t for the high (low) institutional ownership quintile/decile. Turnover (daily) is calculated as shares traded on day t divided by outstanding shares on that day. Consistent with previous literature such as Campbell et al (1993) and Llorente et al. (2002), Turnover proxies for trading volume of individual stocks. LogTurnover(t) is computed as log[turnover(t) + 0.00000255]. A small constant is added to the turnover before taking the log to cater for situations where the trading volume on a particular day may be zero. The value of the constant is chosen to maximize the normality of the distribution of daily trading volume. Detrended V(t) = logTurnover(t) – (average of past 200 days' logTurnover)
The Turnover transformation we use is consistent with previous literature such as Lo and Wang (2001), Llorente et al (2002), and Corvig and Ng(2004). We use two measures of private information PI1 and PI2. PI1 is computed as log[$(1 - r^2)/ r^2$ ] where $r^2$ is obtained from the following regression:

$$R_j = a + bR_m + cR_i + e$$

where $R_j$, $R_m$ and $R_i$ are return on security j, market return, and security j's industry return respectively. For each month, we regress each firm's daily return on the market and 3-digit SIC value-weight industry returns. $(1-r^2)$ is a proxy for private information on the stock. This measure is used in Chen et al. (2007). We use the log transformation of this measure because in some years, $1-r^2$ is leptokurtic and negatively skewed. This transformation is used in Durnev, Morck, and Yeung (2004).

We test whether the coefficient $c_{hi}$ is less than the coefficient $c_{lo}$ in our models. A significant difference validates hypothesis 1.
The two core models for testing hypothesis 2 are models 3 and 4 below:

Model 3: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf|$
: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf|$
Model 4: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf|$
: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf|$

$MV_{(t)}$ is detrended logTurnover of the Market in period t and is transformed in the same way as $V_t$.

|RmRf| is the absolute value of the difference between the return on the market and the risk free rate. $MV_{(t)}$* |RmRf| proxies for publicly available market information and the

coefficient e on the cross product term measures the effect of market information. We follow here Corvig and Ng (2007), Llorente et al. (2002) and Durnev and Nain (2007).

We test whether the coefficient $e_{hi}$ is more than the coefficient $e_{lo}$ in our models. A significant difference validates hypothesis 2.

It is necessary to examine whether our findings are driven by missing variables. For this purpose, we develop eight additional models to perform robustness tests. These models progressively add control variables to our core models – variables that extant literature has shown to significantly influence stock trading volume.

Several researchers have documented the relationship between stock return volatility and trading volume, for example Lamoureux and Lastrapes (1990) and Lee and Rui (2002). These studies provide convincing evidence of a contemporaneous as well as dynamic relationship between return volatility and trading volume. Accordingly, we sequentially introduce stock and market volatility into Models 3 and 4.

Model 5: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}*|RmRf| + f_{hi}*stk\_Volat_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}*|RmRf| + f_{lo}*stk\_Volat_{lo,t}$

Model 6: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}*|RmRf| + f_{hi}*stk\_Volat_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}*|RmRf| + f_{lo}*stk\_Volat_{lo,t}$

Model 7: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}*|RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}*|RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t}$

Model 8: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}*|RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}*|RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t}$

Stock volatility (stk_Volat) is computed as the volatility of the daily stock return over the past thirty days. Similarly market volatility (M_Volat) is computed as the volatility of the market return over the past thirty days.

Researchers such as Gallant et al (1992) find that large price movements are followed by high volume. Accordingly, we introduce stock price run up as a control variable in Models 9 and 10.

Model 9: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}*|RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t} + h_{hi}*stk\_runnup_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}*|RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t} + h_{loi}*stk\_runnup_{lo,t}$

Model 10: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}*|RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t} + h_{hi}*stk\_runnup_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}*|RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t} + h_{loi}*stk\_runnup_{lo,t}$

Stock price run up (stk_runnup) is measured as the thirty day run up in stock price.

There is conflicting evidence on the relationship between momentum and volume. Lee and Swaminathan (2000) show a relationship between turnover and momentum profits. Connolly and Stivers (2003) also evidence such a relationship but Scott et al. (2003) attribute this observed relationship to underreaction to earnings news. They find that the interaction between momentum and volume disappears when a stock's growth rate and earnings-related news are controlled for. It is necessary, therefore, to add momentum as a control variable regardless of whether there is a direct relationship or whether this observed relationship is a proxy for growth rate and the underreaction to earnings news. The addition of momentum generates Models 11 and 12.

Model 11: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| +f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t} + h_{hi}*stk\_runnup_{hi,t} + M\_momentum_{t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t} + h_{loi}*stk\_runnup_{lo,t} + M\_momentum_{t}$

Model 12: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t} + h_{hi}*stk\_runnup_{hi,t} + M\_momentum_{t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t} + h_{loi}*stk\_runnup_{lo,t} + M\_momentum_{t}$

Market momentum (M_momentum) is measured as the raw returns of the market in the past two months.

The above SUR regression models require the formation of high and low institutional ownership portfolios and use portfolio returns in the regressions. In order to test the robustness of our results using individual stock returns , we perform cross-sectional regressions using the first-order autoregressive model of Corvig and Ng (2004) to study the differential impact of private and public information on trading volume.

The basic model takes the following form:

$$V_{i,t+1} = \Sigma\, a_{0,ik}\, D_{k,t+1} + a_{1,i}V_{i,t} + \varepsilon_{i,t+1}$$

Where:

$V_{i,}$ is detrended log turnover, and $D_{k,t+1}$ are the day-of-week dummy variables.

The model is then augmented by adding private information flow. With the two measures of private information we use in this paper, we get the following two versions:

$$V_{i,t+1} = \Sigma\, a_{0,ik}\, D_{k,t+1} + a_{1,i}V_{i,t} + a_{2,i}V_{i,t}PI1_{i,t} + \varepsilon_{i,t+1}$$

And

$$V_{i,t+1} = \Sigma\, a_{0,ik}\, D_{k,t+1} + a_{1,i}V_{i,t} + a_{2,i}V_{i,t}PI2_{i,t} + \varepsilon_{i,t+1}$$

The coefficient $a_{1,I}$ for $V_{i,t}$ represents the constant component of volume autocorrelation whereas the coefficient $a_{2,i}$Vfor $V_{i,t}$PI represents the effect on volume autocorrelation that varies with the flow of information.

## DATA

We start with all the NYSE firms in the CRSP database during the period 1990-2007. We delete financials (6000-6999) and utilities (4000-4949). We also exclude ADRs and REITs, and firms with firm-year observations with less than 30 days of trading. Finally we delete from our sample firms that do not have common shares traded as indicated by CRSP share codes 10 or 11. We conduct two parallel set of tests; in one, we dice the sample into ten groups (deciles) in descending order of institutional ownership and in the other we dice the full sample into five groups (quintiles) based on the same criterion. Stocks in each group constitute an equal weighted portfolio. The portfolios are rebalanced each year. The model inputs come from the top and bottom decile (quintile) portfolios – the two being representative of high and low institutional ownership. We also divide the full sample period into two sub periods 1990-97 and 1998-2007 to test for the stability of the results.

Table 1 gives the descriptive statistics of the high (top decile) and low (bottom decile) institutional ownership portfolios in our sample. In all of the attributes listed in the table, there is a statistically significant difference in the mean and median of the high and low institutional ownership portfolios. It is self-evident from the table that the high institutional ownership portfolio is comprised of firms with larger size, higher stock price (PRC), greater stock turnover, and a larger holding period profit (Hpret). Consistent with the hypotheses we are investigating, both the measures of private information, PI1 and PI2, are larger in low-institutional ownership firms as compared to the high-institutional ownership firms.



## ANALYSIS OF RESULTS

### Core model augmented with standard control variables

Table 2 reports the results of estimating models 1 to 12 with a ten group sort of the data based on decreasing institutional ownership. Table 3 reports the result of a similar sort but into five groups. Hereafter, we refer to the two variables V*PI1 and V*PI2 jointly as Vt.Inf. We focus on the coefficients of Vt.Inf for hypothesis 1, and the coefficient of Mkt.Inf for hypothesis 2. A negative coefficient on Vt.Inf indicates that private information and turnover are negatively related. If the Vt.Inf coefficient is negative for both the high and low institutional deciles/quintiles, then a larger negative coefficient for the high institutional group supports hypothesis 1. If the coefficients are positive for both the high and low institutional groups then a smaller positive for the high institutional group backed by a significant F-value for the model would also support hypothesis 1.

The coefficient for Mkt.Info provides evidence for the veracity of hypothesis 2. A more positive coefficient for the high institutional ownership decile/quintile or a less negative one compared to the low institutional ownership group implies that the less overconfident investors i.e. the high institutional group relies more on public information (Mkt.Info) as compared to the more overconfident investors i.e. the less institutional group. This is supportive of hypothesis 2. Models 1 and 2 for both the decile and quintile sort validate hypothesis 1. The coefficients for Vt.Inf with PI1 are -0.0483 and -0.0193 for the high and low institutional ownership in Model 1 of Table 2, and the F-value is 2.98 which is significant at the 10% level. When PI2 is used as the measure of private information in Model 2, the coefficients for Vt.Inf are -0.0271 and 0.0075 for the high and low specifications with an F-value of 3.94 which is significant at the 5% level. The results in Table 3 with the quintile sort are even more significant. The coefficients for Vt.Inf (PI1) are -0.0766 and 0.0015 for the high and low specifications respectively with an F-value of 23.75 which is significant at the 1% level. With PI2, the results are similar. The coefficients for Vt.Inf with PI2 are -0.4425 and 0.0136 for the high and low regressions respectively and the F-value of 20.31 is significant at the 1% level.

Then we introduce detrended market turnover (MVt) and market information (M.Info) into the models, the latter as a proxy for public information, and test hypotheses 1 and 2 together in Models 3 and 4. In the decile sort, the Vt.Info coefficients for both Models 3 and 4 support hypothesis 1; the high coefficient for Vt.Info (PI1) (-0.0286) is more negative than the low coefficient (-0.0256) in Model 3 and similarly the high coefficient of Vt.Info (PI2) (-0.0219) is smaller than the low coefficient (0.0068) in Model 4, but the F-values for hypotheses 1 is significant only in Model 4. For the public information proxy, the coefficients for M.Info in Models 3 and 4 are supportive of hypothesis 2 i.e. the positive coefficients of the high equations (8.5719 and 9.0401) are greater than the positive coefficients of the low equations (5.3324 and 6.1926) in Models 3 and 4. However the F-values for hypothesis 2 in both models are not significant. With the quintile sort, both Hypotheses 1 and 2 are validated. The signs and sizes of the coefficients for Vt.Info and M.Info are as predicted by hypotheses 1 and 2, and the F-values for the hypotheses are significant.

Next we start introducing the control variables. The first control variable introduced is stock volatility. The estimation results after the introduction of this variable are provided in Models 5 and 6 of Tables 2 and 3. In the decile sort (Table 2), the sign and magnitude of Vt.Info (PI2) coefficients in the high and low equations of Model 6 are supportive of hypothesis 1 (-0.0199<.0038) and the F-value for the hypothesis is significant albeit at the 10% level. But the high and low coefficients of Vt.Info (PI1) do not have the expected signs and sizes in Model 5, and neither is the F-value significant for hypothesis 1. For the public information proxy, MInfo, the sign-size combinations in both Models 5 and 6 are supportive of hypothesis 2 but the F-values of both models are not significant. The t-statistics of the control variable stock volatility are significant for the high equations of Model 5 and 6, but not so for the low equations of the SUR model. Thus stock volatility appears to be significant only for portfolios with high institutional ownership. In the quintile sort (Table 3), the sign-size combinations for the coefficients of Vt.Info with PI1 and PI2, and the public information proxy MVt are supportive of hypotheses 1 and 2 in both Models 5 and 6, but the F-values for both hypotheses in both the models are not significant.

The next control variable introduced is market volatility (M. Volatility). The estimation results after the introduction of this variable are provided in Models 7 and 8 of Tables 2 and 3. In the decile sort (Table 2), the sign and magnitude of the high and low coefficients of Vt.Info (PI2) are supportive of hypotheses 1 (-0.0198<.0077) in Model 8 and the F-value for the hypothesis is significant albeit at the 10% level. But the high and low coefficients of Vt.Info (PI1) do not have the expected signs and sizes in Model 7, and neither is the F-value significant for hypothesis 1. For the public information proxy, M.Info, the sign-size combinations in both Models 7 and 8 are supportive of hypothesis 2 but the F-values of both models are not significant. The t-statistics of the control variable market volatility are significant for the low equations of Model 7 and 8, but not so for the high equations. Thus while stock volatility appears to be significant only for portfolios with high institutional ownership, market volatility appears to be significant only for portfolios with low institutional ownership. In the quintile sort (Table 3), the sign-size combinations for the coefficients of Vt.Info with PI1 and PI2, and the public information proxy

M.Info for both Models 7 and 8 are supportive of hypotheses 1 and 2, but the F-value for only hypothesis 2 of Model 7 is significant while the other F-values are not.

Stock run up (Stk_Runup) is then introduced as a dependent variable. The estimation results after the introduction of this variable are provided in Models 9 and 10 of Tables 2 and 3. In the decile sort (Table 2), the sign and magnitude of the coefficients of Vt.Info with PI2 in the high and low equations of Model 10 are supportive of hypothesis 1 (-0.0177<.0114) and the F-value for the hypothesis is significant albeit at the 10% level. But the high and low coefficients of Vt.Info (PI1) do not have the expected signs and sizes in Model 9, and the F-value for hypothesis 1 is not significant. For the public information proxy, M.Info, the sign-size combinations in both Models 9 and 10 are supportive of hypotheses 2 but the F-values of both models are not significant. The t-statistics of the control variable Stock run up are significant for the low equations of Model 9 and 10, but not so for the high equations. Interestingly, stock market volatility that was previously significant only for the high equations, now becomes significant for the low equations as well but with a negative sign. Thus stock market volatility positively impacts turnover of stocks with high institutional ownership, but it negatively impacts turnover of stocks with low market volatility. In the quintile sort (Table 3), the sign-size combinations for the coefficients of Vt.Info with both PI1 and PI2, and the public information proxy M.Info are supportive of hypotheses 1 and 2 in both Models 9 and 10, and the F-values are significant for hypothesis 1 for Model 9 but not so for Model 10. However, the F-values for hypothesis 2 are significant in both Models 9 and 10.

The last control variable introduced in the model is market momentum (M.Momentum). The estimation results after the introduction of this variable are provided in Models 11 and 12 of Tables 2 and 3. In the decile sort, the sign and magnitude of the coefficients of Vt.Info (PI2) in the high and low equations of Model 12 are supportive of hypothesis 1 (-0.0176<0.0115) and the F-value for the hypothesis is significant albeit at the 10% level. But the high and low coefficients of Vt.Info (PI1) in Model 11 do not have the expected signs and sizes, and the F-value for hypothesis 1 is not significant. For the public information proxy, M.Info, the sign-size combinations in both Models 11 and 12 are supportive of hypotheses 2 but the F-values of both models are not significant. The t-statistics for the newly entered control variable, market momentum, are not significant in any equation of Models 11 and 12. In the quintile sort (Table 3), the sign-size combinations for the coefficients of Vt.Info with PI1 and PI2, and the public information proxy M.Info are supportive of hypotheses 1 and 2 in both Models 11 and 12. The F-values for hypotheses 1 and 2 are significant for Model 11, but not significant for Model 12.

**Estimation results with truncated data**

We further explore the validity of the hypotheses by constraining the data only to certain time periods and/or by excluding firms with certain characteristics from the data set. Thus we estimate the full model i.e. the core model with the control variables for the sub-periods 1990-1997 and 1998-2006. We also study the effect of excluding firms from the data set based on share price, size and liquidity; in each case we drop ten percent of the firms having the lowest share price, size, and liquidity. The results are summarized in Table 4. The first column provides a description of the data set and sort, the second summarizes the result of estimating the model for the whole period 1990-2006. The third and fourth column summarizes the results of estimating the model for the sub-periods 1990-1997 and 1998-2006. The first two rows summarize the results of the full model estimated using a decile and a quintile sort respectively.

The sub-period results indicate that the propositions of hypotheses 1 and 2 had greater validity in the first sub-period 1990-1997 as compared to the second sub-period 1998-2006 because the number of models in which the hypotheses were found significant decline perceptibly in the second sub period.

Rows 3 and 4 summarize the results of estimating the full model on a data set in which ten percent of the firms with the lowest share prices for that year have been excluded. In the decile sort (Row 3), the results for hypothesis 1 for the whole period improve marginally over the full data set in that the hypothesis becomes significant in Model 3 in addition to the models in which it was previously significant. The results for hypothesis 2 do not change. The results for the sub-periods are illuminating. There is a dramatic increase in the significance of hypothesis 1 from being significant only in Model 2 during the first period to being significant in all twelve models in the second period. There is an equally dramatic turnaround in hypothesis 2. In the first sub-period, hypothesis 2 was significant in all models but with the wrong sign, implying that the turnover of stocks with low institutional ownership was more related to public information as compared to stocks with high institutional ownership. However, this position reversed in the second sub-period so that hypothesis 2 registered significance in seven models with the right sign. As a result of this dramatic turnaround, not surprisingly, hypothesis 2 does not show significance for the whole period as the effects of the two sub-periods are opposite and cancel each other out. Overall the results show that the propositions of hypotheses 1 and 2 are strongly valid for the period 1998-2006 once the firms with low value shares are excluded. The final conclusions from the quintile sort (Row 4) are similar – the propositions of hypotheses 1 and 2 are valid, more so for the second sub-period than the first.

Rows 5 and 6 summarize the results of estimating the full model on a data set in which ten percent of the firms with the lowest log size for that year have been excluded. In the decile sort (Row 5), the results are generally poorer than for the full data set - both in the whole period and in the sub-periods. The only exception is hypothesis 2 in the second sub-period which shows significance in Models 3 and 4 whereas it was insignificant in all models in the full data set. Similar to the observation made in data set that excluded firms on size, we find that there is a dramatic reversal in the significance of hypothesis 2 proposition from the first to the second sub-period. In the first sub-period, hypothesis 2 was significant in all models but with the wrong sign. In the second sub-period, all the models had the correct sign and the hypothesis was significant in Models 3 and 4. In the quintile sort (Row 6), the results for the whole period and the second sub-period are not as good as for the full data set. However, the first sub-period is an exception. The results for this sub-period are better in the truncated data set particularly for hypothesis 2 which is significant in all models. The results suggest that overconfident traders use more private information in small stocks and that is the reason why the significance of private information drops when small size stocks are excluded.

Rows 7 and 8 summarize the results of estimating the full model on a data set in which ten percent of the firms with the lowest liquidity for that year have been excluded. Liquidity is measured using the Amihud formula which defines liquidity as absolute return divided by dollar volume. In both the decile and quintile sort, deleting the bottom ten percent liquidity stocks significantly lowers the likelihood that hypothesis 1 is supported in the full period as well as the sub-periods. This implies that overconfident traders use more private information in low-liquidity stocks. The results for hypothesis 2 are more mixed. As compared to the full data set, the results for the significance of hypothesis 2 are marginally better in terms of significance for the whole period, are the same for the first sub-period, and are considerably better for the second sub-period in the decile sort (Row 7). In the quintile sort, the results for hypothesis 2 are better only for the second sub-period.

In conclusion, sub-sample results suggest that overconfident traders use more private information in small and/or low-liquidity stocks. Deleting the bottom ten percent small stocks or the bottom ten percent liquidity stocks in each year significantly lowers the likelihood that hypothesis 1 is supported. This observation applies to both the quintile and decile classifications. Deleting the bottom ten percent of the low-price stocks causes no significant change in the results of hypothesis 1. This suggests that low share price does not encourage the use of private information among overconfident traders.

**Results of the first-order autoregressive model**

In the SUR models described above, the inputs used in the regression are the equal weighted portfolio averages. In the autoregressive model, the regressions are performed using time-series cross-sectional method that employs individual stock data. The results are presented in Table 5. Panel A presents the results of the base model in which the lagged detrended log of the turnover and the day of the week dummy variables are the only explanatory variables. The results show the presence of strong first-order autocorrelation in both the high and low institutional ownership firms. However, the low institutional ownership firms exhibit a higher serial correlation in their trading volume as compared to high institutional firms.

The model in Panel A is then augmented by introducing Vt.Info (PI1) as the measure of private information. The estimation of the resultant model is presented in Panel B. The results show that stocks with lower institutional ownership have a higher serial correlation in trading volume and are more influenced by private information as measured by PI1.

Panel C model is similar to the Panel B model except that the private information measure used is PI2. The results show that stocks with lower institutional ownership have a higher serial correlation in trading volume and the volume is unaffected by private information as measured by Vt.Info.

Overall the results in Table 5 suggest that stocks with low-institutional ownership exhibit significantly higher serial correlation in trading volume. There is some evidence, as seen in Panel B, that the trading volume of stocks of lower-institutional ownership is affected by traders' private information (PI1).

**CONCLUSION**

In this paper, we add to the overconfidence literature by accounting for the source of information and the type of investor. The literature suggests that overconfidence is a major determinant of stock trading volume. Trading is triggered by the arrival of new information which may be public or private. Overconfident investors overestimate the precision of their private information signals and trade more than is warranted by the incoming signal. We postulate that private investors are more prone to overconfidence bias as compared to institutional investors. This implies that turnover in firms with low institutional ownership will be driven more by private information (hypothesis 1) while turnover in firms with high institutional ownership will be driven more by public information (hypothesis 2). This is the essence of the two hypotheses we explore. We use two measures of private information, PI1 and PI2, and sort the firms on the basis of institutional ownership into groups. We use both a ten group sort and a five group sort and employ SUR models to study the hypotheses by examining the differences between the group with the highest and lowest institutional ownership. We find strong evidence in favor of hypothesis 1. PI1 and PI2 are found to be significant both in the decile and quintile sort when either PI1 or PI2 is the only explanatory variable besides the lagged value of the independent variable, turnover. Even in the full models with all the control variables, PI2 is significant in the decile sort (Model 12) and PI1 in the quintile sort (Model 11). In particular, PI2 is significant in all six models of the decile sort that use PI2 as a measure of private information. Evidence in favor of hypothesis 2 is more mixed. $MV_{(t)}$* |RmRf| proxies for publicly available market information. Without the addition of control variables, the proxy is significant in the decile sort when PI2 is used as a measure of private information (Model 4). In the quintile sort, the public information proxy is significant regardless of whether PI1 or PI2 is used as a measure of private information. In the full model with all the control variables, the public information proxy is significant only in the quintile sort with PI1 as a measure of private information (Model 11). However, the sub-period analysis with truncated data finds strong evidence in support of hypothesis 2 in the second sub-period 1998-2006 when ten percent of the stocks having the lowest liquidity or lowest price are deleted from the data. This implies that managers of high institutional ownership stocks in the most recent period do rely more on public information for their stocks but not so if the stocks have a low price or the firms have poor liquidity. Finally, the first order autoregressive model provides further support for hypothesis 1. The results show that turnover volume of stocks with lower institutional holdings is affected by trader's private information (PI1).

## REFERENCES

Barber, B. and T. Odean, 2000, Trading is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors, *Journal of Finance*, Vol. LV, No. 2, 773-806.

Barber, B. and T. Odean, 2008, All that Glitters: The Effect of Attention and News on the Buying Behavior of Individuals and Institutional Investors, *The Review of Financial Studies*, Volume 21 No. 2, 785-818.

Benos, A.V. 1998, Aggressiveness and Survival of Overconfident Traders, *Journal of Financial Markets*, 1, 353-383.

Black, F. 1986, Noise, *Journal of Finance*, Vol. XLI, No. 3, 529-543.

Campbell, J.Y., S.J. Grossman and J. Wang, 1993, Trading Volume and Serial Correlation in Stock Returns, *Quarterly Journal of Economics*, 108, 905-939.

Chae, J., 2005. Trading Volume, Information Asymmetry, and Timing Information, *Journal of Finance*, Vol LX, No. 1, 413-442.

Chen, Q., I. Goldstein, and W. Jiang, 2007, Price Informativeness and Investment Sensitivity to Stock Price, The Review of Financial Studies, Volume 20 No. 3, 619-650.

Cho, J. and J. Kyoosung, 2006, Overconfidence and Information: The Differences Between Individuals and Institutions, working paper (www.ssm.com).

Chuang, W.I, and B-S. Lee, 2006, An Empirical Evaluation of the Overconfidence Hypothesis, *Journal of Banking and Finance*, 30, 2489-2491.

Connoly, R. and C. Stivers, 2003, Momentum and Reversals in Equity-Index Returns During Periods of Abnormal Turnover And Return Dispersion, *Journal of Finance*, Vol. 58, 2003, pp 1521-1556.

Covrig, Vincentiu, and Lilian Ng, 2004, Volume autocorrelation, information, and investor trading, *Journal of Banking & Finance*, 28, 2155-2174.

Daniel, K., D. Hirshleifer and A. Subrahmanyam, 1998, A Theory of Overconfidence, Self-Attribution, and Security Market Under- and Over-Reactions, *Journal of Finance*, 53, 1839-1886.

De Bondt, W.F. and R.H. Thaler, 1995, Financial Decision Making in Markets and Firms: A Behavioral Perspective, in *Handbooks in Operations Research and Management Sciences*, 9, *Finance*, R.A. Jarrow, V. Maksimovic and W.T. Ziemba (editors), Amsterdam, Elsevier, 385-410.

Dow, J. and G. Gorton, 2006, Noise Traders, NBER Working Paper Series, Working Paper 12256.

Durnev, A., R. Morck and B. Yeung, 2004, Value Enhancing Capital Budgeting and Firm-Specific Stock Variation," *Journal of Finance*, 59, 65-105.

Durnev, A., and A. Nain, 2007, Does Insider Trading Regulation Deter Private Information Trading? International Evidence. Working Paper. Available at SSRN: http://ssrn.com/abstract=950105

Easley, D., N. Kiefer and M. O'Hara, 1996, Cream Skimming or Profit Sharing? The Curious Role of Purchased Order Flow, *Journal of Finance*, 51, 811-833.

Gallant, A.R., P.E. Rossi and G. Tauchen, 1992, Stock Price and Volume, *The Review of Financial Studies*, Volume 5 No. 2 pp 199-242.

Gervais, S. and T. Odean, 2001, Learning To Be Overconfident, *Review of Financial Studies*, 14, 1-27.

Glaser, Markus, Markus Noth, and Martin Weber, 2004, Behavioral Finance in *Blackwell Handbook of Judgment and Decision Making*, D.J. Koehler and N. Harvey (Editors), Malden, Mass.: Blackwell, 527-546.

Glaser, Markus, and Martin Weber, 2007, Overconfidence and Trading Volume, *Geneva Risk Insurance Review*, 32, 1-36.

Graham, J.R., and C.R. Harvey, 2002, Expectations of Equity Risk Premia, Volatility, and Asymmetry, Working Paper, Duke University.

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

Hilton, D.J., 2001, The Psychology of Financial Decision Making: Application to Trading, Dealing and Investment Analysis, *Journal of Psychology and Financial Markets*, 2, 37-53.

Kirchler, E. and B. Maciejovsky, 2002, Simultaneous Over- and Under-Confidence: Evidence From Experimental Asset Markets, *Journal of Risk and Uncertainty*, 25, 65-85.

Lamoureux, C.G. and W.D. Lastrapes. 1990, Heteroscedasticity in stock return data: Volume versus GARCH effects. *Journal of Finance* 45 (1): 221-229.

Lee, B-S, O.M. Rui, 2002, The dynamic relationship between stock returns and trading volume: Domestic and cross-country evidence, *Journal of Banking and Finance* 26 (1): 51-78.

Lee, C.M.C. and B. Swaminathan, 2000, "Price momentum and trading volume," *Journal of Finance*, Vol. 55, 2017-2069.

Llorente, G., R. Michaely, G. Saar, and J. Wang, 2002, Dynamic volume-return relationship of individual stocks, *Review of Financial Studies* 15, 1005-1047.

Morck, R., B. Yeung and W. Yu, 2000, The Information Content of Stock Markets: Why Do Emerging Markets Have Synchronous Stock Price Movements? *Journal of Financial Economics*, 59, 215-260.
Odean, T., 1998, Volume, Volatility, Price, and Profit When All Traders Are Above Average, Journal of Finance, 53, 1887-1934.
Odean, T., 1999, Do Investors Trade Too Much? *American Economic Review*, 89, 1279-1298.
Roll, R., 1988, $R^2$, *Journal of Finance*, 43, 541-566.
Scott, J., M. Stumpp and P. Xu, 2003, "News, Not Trading Volume Builds Momentum," *Financial Analyst Journal*, Vol. 59, pp 45-54.
Shefrin, H., 2005, A Behavioral Approach to Asset Pricing, Elsevier Academic Press, 131.
Statman, Meir, Steven Thorley, Keith Vorkink, 2006, Investor Overconfidence and Trading Volume, *The Review of Financial Studies*, 19, 1531-1565.
Wang, F.A. 1998, Strategic Trading, Asymmetric Information and Heterogeneous Prior Beliefs, *Journal of Financial Markets*, 1, 321-352.

Table 1: Descriptive statistics of share price, daily volume, firm size, and private information measures

| | High-Institutional Ownership (top decile) | | | Low-Institutional Ownership (bottom decile) | | | High - Low | High - Low |
|---|---|---|---|---|---|---|---|---|
| | Mean | Median | Std Dev. | Mean | Median | Std Dev. | Mean (t-value) | Median (p-vale) |
| PRC | 32.30 | 28.25 | 26.71 | 5.32 | 2.25 | 14.57 | 746.8 | <0.0001 |
| Log(size) | 20.71 | 20.75 | 1.31 | 17.49 | 17.16 | 1.90 | 1147.3 | <0.0001 |
| Turnover | 0.0068 | 0.0042 | 0.0112 | 0.0028 | 0.0007 | 0.0145 | 186.2 | <0.0001 |
| Detrended Turnover | 0.091 | 0.075 | 1.089 | -0.016 | 0.063 | 2.530 | 78.1 | <0.0001 |
| Hpret | 0.224 | 0.139 | 0.544 | 0.085 | -0.087 | 1.23 | 88.5 | <0.0001 |
| PI1 | 0.534 | 0.593 | 1.930 | 2.29 | 2.24 | 1.62 | 577.6 | <0.0001 |
| PI2 | 0.576 | 0.626 | 0.296 | 0.80 | 0.89 | 0.25 | 493.1 | <0.0001 |
| No. of observations | 734699 | | | 702120 | | | | |

PRC is share price; Hpret is holding period return and PI1 and PI2 are measures of private information as defined in the text.

Table 2
Results of the SUR models when run for the full sample period on the highest and lowest decile portfolios sorted on the basis of institutional ownership

Private information measure is PI1 in odd numbered models and PI2 in even numbered models

| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | | Model 6 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | High | Low | High | Low | High | Low | High | Low | High | Low | High | Low |
| intercept | 0.0341 | 0.0146 | 0.0337 | 0.0158 | 0.0293 | 0.0243 | 0.0287 | 0.0234 | -0.0075 | 0.0261 | -0.0062 | 0.0218 |
| | 11.21 | 2.39 | 11.31 | 2.63 | 9.37 | 3.94 | 9.38 | 3.97 | -0.95 | 1.96 | -0.81 | 1.67 |
| Vt | 0.6349 | 0.818 | 0.6071 | 0.7368 | 0.7408 | 0.8473 | 0.7279 | 0.7823 | 0.7279 | 0.8411 | 0.7763 | 0.7828 |
| | 31.52 | 40.38 | 36.91 | 56.07 | 24.41 | 41.22 | 26.73 | 57.58 | 23.79 | 41.31 | 26.64 | 57.59 |
| Vt.Info | -0.0483 | -0.0193 | -0.0271 | 0.0075 | -0.0286 | -0.0256 | -0.0219 | 0.0068 | -0.0139 | -0.0274 | -0.0199 | 0.0038 |
| | -3.31 | -2.28 | -2.28 | 0.66 | -1.92 | -2.03 | -1.67 | 0.61 | -0.93 | -2.34 | -1.52 | 0.52 |
| MVt | | | | | -0.2674 | -0.3129 | -0.2615 | -0.0323 | -0.2667 | -31.28 | -0.2611 | -0.2998 |
| | | | | | -8.52 | -8.34 | -8.55 | -8.71 | -8.61 | -8.34 | -8.59 | -8.16 |
| M.Info | | | | | 8.5719 | 5.3324 | 9.0401 | 6.1926 | 7.1228 | 5.3509 | 7.3799 | 6.9011 |
| | | | | | 6.32 | 1.91 | 6.93 | 2.25 | 5.16 | 1.89 | 5.5 | 2.19 |
| Stk.Volatility | | | | | | | | | 4.3081 | -0.1531 | 3.8377 | 0.2527 |
| | | | | | | | | | 5.81 | -0.13 | 5.05 | 0.21 |
| Mkt.Volatility | | | | | | | | | | | | |
| Stk_Runup | | | | | | | | | | | | |
| M.Momentum | | | | | | | | | | | | |
| F-value Hypothesis 1 | 2.98 | | 3.94 | | 0.03 | | 2.80 | | 0.60 | | 2.67 | |
| | (0.0845) | | (0.0470) | | (0.8588) | | (0.0941) | | (0.4385) | | (0.1000) | |

| | | | | |
|---|---|---|---|---|
| Hypothesis 2 | 1.36 | 1.12 | 0.40 | 0.22 |
| | (0.2428) | (0.2904) | (0.5268) | (0.6383) |

Table 2 (Continued)
Results of the SUR models when run for the full sample period on the highest and lowest decile portfolios sorted on the basis of institutional ownership

| | Model 7 | | Model 8 | | Model 9 | | Model 10 | | Model 11 | | Model 12 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | High | Low | High | Low | High | Low | High | Low | High | Low | High | Low |
| | - | | - | | - | - | - | - | - | - | - | - |
| intercept | 0.0107 | 0.0071 | 0.0094 | 0.0042 | 0.0186 | 0.0189 | 0.0151 | 0.0053 | 0.0168 | 0.0081 | 0.0145 | 0.0052 |
| | -1.35 | 0.44 | -1.22 | 0.27 | -2.14 | -0.64 | -1.81 | -0.33 | -1.94 | -0.49 | -1.74 | -0.33 |
| Vt | 0.6267 | 0.8465 | 0.7237 | 0.7811 | 0.6771 | 0.7943 | 0.6929 | 0.7506 | 0.6749 | 0.7827 | 0.6928 | 0.7506 |
| | 23.6 | 40.94 | 26.56 | 57.42 | 21.14 | 36.04 | 24.91 | 53.35 | 21.09 | 35.37 | 24.9 | 53.53 |
| | - | - | - | | | - | - | | | - | - | |
| Vt.Info | 0.0135 | 0.0254 | 0.0198 | 0.0077 | 0.0041 | 0.0084 | 0.0177 | 0.0114 | 0.0047 | 0.0084 | 0.0176 | 0.0115 |
| | -0.91 | -1.98 | -1.51 | 0.78 | 0.26 | -0.96 | -1.36 | 1.03 | 0.31 | -0.96 | -1.35 | 1.03 |
| | - | - | - | - | - | - | - | - | - | - | - | - |
| MVt | 0.2649 | 0.3317 | 0.2594 | 0.2982 | 0.2451 | 0.2851 | 0.2361 | 0.2701 | 0.2438 | 0.2824 | 0.2374 | 0.2303 |
| | -8.55 | -8.31 | -8.53 | -8.51 | -7.83 | -7.53 | -7.71 | -7.33 | -7.79 | -7.49 | -7.72 | -7.33 |
| M.Info | 7.1241 | 5.9724 | 7.3422 | 4.7011 | 7.4963 | 3.9386 | 7.4432 | 4.6037 | 7.3458 | 4.1917 | 7.2898 | 4.5595 |
| | 5.12 | 1.37 | 5.34 | 1.65 | 5.39 | 1.34 | 5.52 | 1.62 | 5.27 | 1.45 | 5.04 | 1.06 |
| | | - | | - | | - | | - | | - | | - |
| Stk.Volatility | 6.3733 | 2.5446 | 5.8236 | 2.0748 | 7.3659 | 3.9763 | 6.6811 | 3.1629 | 7.3159 | 3.7403 | 6.5519 | 3.1688 |
| | 2.96 | -1.6 | 2.75 | -1.32 | 3.43 | -2.5 | 3.15 | -2.02 | 3.39 | -2.36 | 3.09 | -2.03 |
| | - | | - | | - | | - | | - | | - | |
| M.Volatility | 2.2394 | 5.3165 | 1.8063 | 5.5074 | 2.7342 | 8.5666 | 2.4044 | 7.5618 | 2.7581 | 8.5534 | 2.2693 | 7.5754 |
| | -1.01 | 2.27 | -0.84 | 2.24 | -1.22 | 3.54 | -1.09 | 3.28 | -1.23 | 3.85 | -1.02 | 3.29 |
| Stk_Runup | | | | | 0.1098 | 0.3814 | 0.1208 | -3108 | 0.0144 | 0.3309 | 0.1153 | 0.3103 |
| | | | | | 1.71 | 2.7 | 1.89 | 2.84 | 1.62 | 2.98 | 1.81 | 2.83 |
| M.Momentum | | | | | | | | | - | - | - | - |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | 0.6422 0.2051 | 0.6429 0.1839 |
| | | | | | -1.64 -0.25 | -1.64 -0.22 |
| F-value | | | | | | |
| Hypothesis 1 | 0.47 | 2.67 | 0.52 | 2.90 | 0.56 | 2.87 |
| | (0.4934) | (0.1000) | (0.4495) | (0.0887) | (0.4565) | (0.0902) |
| Hypothesis 2 | 1.21 | 0.88 | 1.31 | 1.04 | 1.23 | 0.96 |
| | (0.2708) | (0.3483) | (0.2519) | (0.3086) | (0.2674) | (0.3269) |

The table shows SUR estimates of the models:

Model 1: $V_{hi,t+1} = a_{hi} + b_{hi}*V_{hi,t} + c_{hi}*V_{hi,t}*PI1_{hi,t}$

: $V_{lo,t+1} = a_{lo} + b_{(lo)}*V_{lo,t} + c_{lo}*V_{lo,t}*PI1_{lo,t}$

Model 2: $V_{hi,t+1} = a_{hi} + b_{hi}*V_{hi,t} + c_{hi}*V_{hi,t}*PI2_{hi,t}$

: $V_{lo,t+1} = a_{lo} + b_{(lo)}*V_{lo,t} + c_{lo}*V_{lo,t}*PI2_{lo,t}$

Model 3: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf|$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf|$

Model 4: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf|$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf|$

Model 5: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t}$

Model 6: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t}$

Model 7: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t}$

Model 8: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t}$

Model 9: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t} + h_{hi}*stk\_runnup_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t} + h_{loi}*stk\_runnup_{lo,t}$

Model 10: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t} + h_{hi}*stk\_runnup_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t} + h_{loi}*stk\_runnup_{lo,t}$

Model 11: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t} + h_{hi}*stk\_runnup_{hi,t} + M\_momentum_{t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| +$

$f_{lo}$*stk_Volat$_{lo,t}$ + $g_{lo}$*M_Volat$_t$ + $h_{loi}$*stk_runnup$_{lo,t}$ + M_momentum$_t$

Model 12: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| +$
$f_{hi}$*stk_Volat$_{hi,t}$ + $g_{hi}$*M_Volat$_t$ + $h_{hi}$*stk_runnup$_{hi,t}$ + M_momentum$_t$
: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| +$
$f_{lo}$*stk_Volat$_{lo,t}$ + $g_{lo}$*M_Volat$_t$ + $h_{loi}$*stk_runnup$_{lo,t}$ + M_momentum$_t$

The subscripts hi and lo denote the portfolio with the highest and lowest institutional ownership respectively and t is a subscript that tracks the periods. V is detrended log turnover, PI1 and PI2 are two measures of private information, V*PI1 (PI2) is V.Info; MV is detrended log turnover of the market, and |RmRf| is the absolute value of the difference between the return on the market and the risk free rate. $MV_{(t)}$* |RmRf| is M.Info and proxies for publicly available market information; Stk.Volatility is stock volatility and M_Volat is market volatility; stk_runnup is stock runup and M_momentum is market momentum.

Table 3
Results of the SUR models when run for the full sample period on the highest and lowest quintile portfolios sorted on the basis of institutional ownership

Private information measure is PI1 in odd numbered models and PI2 in even numbered models

| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | | Model 6 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | High | Low | High | Low | High | Low | High | Low | High | Low | High | Low |
| intercept | 0.0234 | 0.0077 | 0.0222 | 0.0087 | 0.0232 | 0.0183 | 0.0232 | 0.0183 | -0.0065 | 0.0283 | -0.0093 | 0.0298 |
| | 8.31 | 1.94 | 8.09 | 2.21 | 8.08 | 4.51 | 7.82 | 4.54 | -0.95 | 3.46 | -1.37 | 3.54 |
| Vt | 0.6919 | 0.8206 | 0.9311 | 0.7899 | 0.7898 | 0.8585 | 0.9479 | 0.8966 | 0.7562 | 0.8093 | 0.8534 | 0.9002 |
| | 35.18 | 38.41 | 15.04 | 13.49 | 21.01 | 46.67 | 12.98 | 15.44 | 19.96 | 40.72 | 11.48 | 14.55 |
| Vt.Info | -0.0766 | 0.0015 | -0.4425 | 0.0136 | -0.0533 | -0.0094 | -0.2929 | -0.0753 | -0.0396 | -0.0017 | -0.1867 | -0.0834 |
| | -5.31 | 0.18 | -5.13 | 0.21 | -3.56 | -1.16 | -3.35 | -1.44 | -2.43 | -1.45 | -2.11 | -1.29 |
| MVt | | | | | -0.2891 | -0.3037 | -0.2799 | -0.2988 | -0.2696 | -0.3007 | -0.2669 | -0.2951 |
| | | | | | -7.81 | -12.12 | -7.81 | -12.14 | -7.48 | -12.09 | -7.46 | -12.01 |
| M.Info | | | | | 8.4841 | 4.5111 | 8.2496 | 4.5272 | 7.2377 | 4.9213 | 7.0517 | 4.9956 |
| | | | | | 6.55 | 2.49 | 6.29 | 2.51 | 5.52 | 2.69 | 5.33 | 2.71 |
| Stk.Volatility | | | | | | | | | 3.4546 | -1.0131 | 3.6761 | -1.2792 |
| | | | | | | | | | 0.81 | -1.41 | 5.31 | -1.49 |
| Mkt.Volatility | | | | | | | | | | | | |
| Stk_Runup | | | | | | | | | | | | |
| M.Momentum | | | | | | | | | | | | |

| F-value | | | | | | |
|---|---|---|---|---|---|---|
| Hypothesis 1 | 23.75 | 20.31 | 7.02 | 4.36 | 2.24 | 0.94 |
| | (0.0001) | (.0001) | (0.0081) | (0.0368) | (0.1345) | (0.3314) |
| Hypothesis 2 | | | 5.65 | 4.87 | 1.92 | 1.44 |
| | | | (0.0175) | (0.0274) | (0.1664) | (0.2301) |

Table 3 (Continued)
Results of the SUR models when run for the full sample period on the highest and lowest quintile portfolios sorted on the basis of institutional ownership

| | Model 7 | | Model 8 | | Model 9 | | Model 10 | | Model 11 | | Model 12 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | High | Low | High | Low | High | Low | High | Low | High | Low | High | Low |
| intercept | -0.0108 | 0.0163 | -0.0107 | 0.0169 | -0.0191 | 0.0079 | -0.0214 | 0.0097 | -0.0185 | 0.0078 | -0.0209 | 0.0097 |
| | -1.53 | 1.63 | -1.94 | 1.63 | -2.44 | 0.74 | -2.77 | 0.93 | -2.38 | 0.74 | -0.271 | 0.92 |
| Vt | 0.7473 | 0.8549 | 0.8449 | 0.9039 | 0.7109 | 0.8029 | 0.7674 | 0.8353 | 0.8098 | 0.8031 | 0.7661 | 0.8364 |
| | 19.59 | 40.72 | 11.34 | 15.56 | 18.34 | 35.72 | 10.14 | 14.02 | 18.03 | 35.72 | 10.13 | 14.26 |
| Vt.Info | -0.0362 | -0.0113 | -0.1848 | 0.0868 | -0.0235 | 0.0019 | -0.1084 | -0.0368 | -0.0234 | 0.0018 | -0.1081 | -0.0368 |
| | -2.38 | -1.29 | -2.07 | -1.34 | -1.65 | 0.22 | -1.21 | -0.57 | -1.53 | 0.22 | -1.21 | -0.57 |
| MVt | -0.2622 | -0.3013 | -0.2602 | -0.2957 | -0.2476 | -0.2775 | -0.2459 | -0.2719 | -0.2496 | -0.2772 | -0.2552 | -0.2781 |
| | -7.25 | -12.12 | -7.24 | -12.03 | -6.81 | -11.08 | -6.81 | -10.99 | 6.79 | -11.08 | -0.79 | -10.98 |
| M.Info | 7.1239 | 4.1145 | 6.8307 | 4.1355 | 7.4629 | 4.4165 | 7.2405 | 4.2867 | 7.3563 | 4.4277 | 7.1388 | 4.3206 |
| | 5.37 | 2.19 | 5.15 | 2.22 | 5.62 | 2.37 | 5.41 | 2.29 | 5.54 | 2.37 | 5.32 | 2.31 |
| Stk.Volatility | 5.9721 | -3.9209 | 5.4626 | -3.2003 | 6.1708 | 3.6623 | 6.2633 | -3.7558 | 6.5816 | -3.6817 | 6.1031 | -3.7786 |
| | 2.59 | -2.57 | 2.33 | -2.67 | 2.91 | -3.04 | 2.73 | -3.12 | 2.84 | -3.95 | 2.66 | -3.15 |
| Mkt.Volatility | -2.1777 | 3.4377 | -1.3525 | 3.5888 | -2.3055 | 4.8508 | -1.6375 | 4.2877 | -2.1905 | 4.8655 | -1.4741 | 4.7983 |
| | -0.92 | 2.19 | -0.57 | 2.28 | -0.89 | 3.66 | -0.69 | 3.03 | -0.92 | 3.07 | -0.62 | 3.04 |
| Stk_Runup | | | | | 0.1494 | 0.2583 | 0.1371 | 0.2842 | 0.1443 | 0.2538 | 0.1316 | 0.2403 |
| | | | | | 2.38 | 3.07 | 2.19 | 2.93 | 2.29 | 3.06 | 2.11 | 2.93 |

| | | | | | | |
|---|---|---|---|---|---|---|
| M.Momentum | | | | | -0.4094 0.0559 | -0.4181 0.1432 |
| | | | | | -1.09 0.11 | 1.11 0.26 |
| F-value | | | | | | |
| Hypothesis 1 | 2.19 | 0.85 | 2.69 | 0.45 | 2.74 | 0.45 |
| | (0.1393) | (0.3562) | (0.0984) | (0.5002) | (0.0988) | (0.5031) |
| Hypothesis 2 | 3.07 | 2.47 | 3.20 | 2.96 | 2.96 | 2.66 |
| | (0.0801) | (0.1160) | (0.0735) | (0.0854) | (0.0855) | (0.1010) |

The table shows SUR estimates of the models:

Model 1: $V_{hi,t+1} = a_{hi} + b_{hi}*V_{hi,t} + c_{hi}*V_{hi,t}*PI1_{hi,t}$

: $V_{lo,t+1} = a_{lo} + b_{(lo)}*V_{lo,t} + c_{lo}*V_{lo,t}*PI1_{lo,t}$

Model 2: $V_{hi,t+1} = a_{hi} + b_{hi}*V_{hi,t} + c_{hi}*V_{hi,t}*PI2_{hi,t}$

: $V_{lo,t+1} = a_{lo} + b_{(lo)}*V_{lo,t} + c_{lo}*V_{lo,t}*PI2_{lo,t}$

Model 3: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf|$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf|$

Model 4: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf|$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf|$

Model 5: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t}$

Model 6: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t}$

Model 7: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t}$

Model 8: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t}$

Model 9: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t} + h_{hi}*stk\_runnup_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t} + h_{loi}*stk\_runnup_{lo,t}$

Model 10: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t} + h_{hi}*stk\_runnup_{hi,t}$

: $V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t} + h_{loi}*stk\_runnup_{lo,t}$

Model 11: $V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI1_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t} + h_{hi}*stk\_runnup_{hi,t} + M\_momentum_{t}$

$$: V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI1_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t} + h_{loi}*stk\_runnup_{lo,t} + M\_momentum_{t}$$

Model 12: $$V_{hi,t+1} = a_{hi} + b_{(hi)}*V_{(hi,t)} + c_{(hi)}*V_{(hi,t)}*PI2_{(hi,t)} + d_{(hi)}*MV_{(t)} + e_{(hi)}*MV_{(t)}* |RmRf| + f_{hi}*stk\_Volat_{hi,t} + g_{hi}*M\_Volat_{t} + h_{hi}*stk\_runnup_{hi,t} + M\_momentum_{t}$$

$$: V_{(lo,t+1)} = a_{(lo)} + b_{(lo)}*V_{(lo,t)} + c_{(lo)}*V_{(lo,t)}*PI2_{(lo,t)} + d_{(lo)}*MV_{(t)} + e_{(lo)}*MV_{(t)}* |RmRf| + f_{lo}*stk\_Volat_{lo,t} + g_{lo}*M\_Volat_{t} + h_{loi}*stk\_runnup_{lo,t} + M\_momentum_{t}$$

The subscripts hi and lo denote the portfolio with the highest and lowest institutional ownership respectively and t is a subscript that tracks the periods. V is detrended log turnover, PI1 and PI2 are two measures of private information, V*PI1 (PI2) is V.Info; MV is detrended log turnover of the market, and |RmRf| is the absolute value of the difference between the return on the market and the risk free rate. $MV_{(t)}*$ |RmRf| is M.Info and proxies for publicly available market information; Stk.Volatility is stock volatility and M_Volat is market volatility; stk_runnup is stock runup and M_momentum is market momentum.

Table 4
Summary of sub-period and truncated data analysis

| Sort | Whole Period | 1990-1997 | 1998-2006 |
|---|---|---|---|
| Decile Daily | H1: significant in models 1,2,4,6,8,10,12<br><br>H2: all insignificant | H1: significant in models 2,4,6,10,12<br><br>H2: all insignificant | H1: significant in 2<br><br>H2: all insignificant |
| Quintile Daily | H1: significant in models 1,2,3,4,9,11<br><br>H2: correct sign in all. Significant in models 3,4,7,9,10,11 | H1: significant in models 1,2<br><br>H2: correct sign in all. Significant in models 3,4,9,10,11,12 | H1: significant in models 1,2,3<br><br>H2: all insignificant |
| Decile Daily No PRC | H1: significant in models 1,2,3,4,6,8,10,12<br><br>H2: all insignificant | H1: significant in model 2<br><br>H2: wrong sign and significant in all models | H1: significant in all models<br><br>H2: correct sign in all. Significant in models 3,4,5,6,7,9,11 |
| Quintile Daily No PRC | H1: significant in models 1,2,3,9,10,11,12<br><br>H2: correct sign in all. All significant. | H1: significant in models 1,2,5,9,10,11,12<br><br>H2: correct sign in all. Significant in models 4,5,6,7,8,9,10,11,12 | H1: significant in models 1,2,3,4,5,6,7,8,9,11<br><br>H2: correct sign in all. Significant in all models |
| Decile Daily No Size | H1: significant in models 1,2,10,12<br><br>H2: all insignificant | H1: significant in models 2,11,12<br><br>H2: wrong sign and significant in all models | H1: all insignificant<br><br>H2: correct sign in all. Significant in models 3,4 |
| Quintile Daily No Size | H1: significant in model 1<br><br>H2: all insignificant | H1: significant in models 1,2,3,8,10<br><br>H2: correct sign in all. Significant in all models. | H1: all insignificant<br><br>H2: all insignificant |
| Decile Daily No Liq | H1: all insignificant | H1: all insignificant | H1: all |

| | | | |
|---|---|---|---|
| | H2: all insignificant except model 3 | H2: all insignificant | insignificant<br><br>H2: correct sign in all. Significant in all models. |
| Quintile Daily No Liq | H1: significant in models 1,3<br><br>H2: correct sign in all. Significant in models 3, 4 | H1: significant in models 1,3,5<br><br>H2: correct sign in all. Significant in models 10, 12 | H1: significant in model 1<br><br>H2: correct sign in all. Significant in models 1,3, 7 |

Table 4 (Continued):

Significant means significant at the one, five or ten percent level.
No PRC means 10% of the firms with the lowest share price have been deleted from the full data set.
No Size means 10% of the firms with the smallest size have been deleted from the full data set.
No Liq means 10% of the firms with the lowest liquidity have been deleted from the full data set.
H1 is hypothesis 1 and H2 is hypothesis 2.

Table 5
Serial correlation and the effect of information flow on stock trading of high versus low institutional ownership (t-values in brackets).

Table 5 Panel A

| | Intercept | $a_1$ | Model Fitness |
|---|---|---|---|
| High | 0.0901 | 0.0099 | F=192.79 |
| | (8.14) | (4.52) | (p<0.0001) |
| Low | -0.165 | 0.042 | F=1127.5 |
| | (-2.39) | (6.48) | (p<0.0001) |
| High-Low | | -0.032 | |
| | | (-4.63) | |

The table shows estimates of the time-series cross-sectional regression on all the individual stocks:

Model : $V_{i,t+1} = \Sigma\, a_{0,ik}\, D_{k,t+1} + a_{1,i}V_{i,t} + \varepsilon_{i,t+1}$
V is detrended log(turnover) and D is day of the week dummy variable.

Table 5 Panel B:

| | Intercept | $a_1$ | a2 | Model Fitness |
|---|---|---|---|---|
| High | 0.0896 | 0.0097 | 0.0007 | F=106.29 |
| | (7.30) | (6.01) | (0.90) | (p<0.0001) |
| Low | -0.1635 | 0.0364 | 0.0017 | F =523.1 |
| | (-0.06) | (5.01) | (1.89)*** | (p<0.0001) |
| High-Low | | -0.032 | -0.0010 | |
| | | (-3.29)) | (-0.67) | |

The table shows estimates of the time-series cross-sectional regression on all the individual stocks:

Model : $V_{i,t+1} = \Sigma\, a_{0,ik}\, D_{k,t+1} + a_{1,i}V_{i,t} + a_{2,i}V_{i,t}PI1_{i,t} + \varepsilon_{i,t+1}$

V is detrended log(turnover), D is day of the week dummy variable and PI1 is a measure of private information.

Table 5 Panel C:

| | Intercept | $a_1$ | a2 | Model Fitness |
|---|---|---|---|---|
| High | 0.0901 | 0.0079 | 0.0011 | F=98.90 |
| | (7.41) | (3.14) | (1.33) | (p<0.0001) |

| | | | | |
|---|---|---|---|---|
| Low | -0.1601<br>(-2.22) | 0.0420<br>(3.48) | 0.0007<br>(1.10) | F =563.9<br>(p<0.0001) |
| High-Low | | -0.034<br>(-2.17)) | 0.0004<br>(0.14) | |

The table shows estimates of the time-series cross-sectional regression on all the individual stocks:

Model : $V_{i,t+1} = \Sigma\, a_{0,ik}\, D_{k,t+1} + a_{1,i}V_{i,t} + a_{2,i}V_{i,t}PI2_{i,t} + \varepsilon_{i,t+1}$

V is detrended log(turnover), D is day of the week dummy variable and PI2 is a measure of private information.