Informativeness of Orders in Electronic Limit Order Book Markets: A Revealed Preference Framework

Sudip Gupta*Vikas Raman†Pradeep Yadav‡

Preliminary and Incomplete, Donot Copy or Post§

May 15, 2019

Abstract

The existence and behavior of "informed traders" plays a foundational role in finance theory, but empirical identification of informed traders has been, at best, challenging. We develop a new measure of trader informativeness in the context of the widespread proliferation of electronic limit order book trading. Our (semi-parametric) measure imposes minimal assumptions on the underlying distribution of private information and the trader's order submission or trading behavior. We show that our measure significantly dominates extant approaches in out of sample tests. We then use our measure to test various hypotheses related to informed trader choice of limit, market and hidden orders in normal, turbulent, and information-intensive periods.

1 Introduction and Motivation

The vast majority of securities are now traded in electronic limit order book markets, particularly equities and particularly outside the US. Even the historical bastions of dealer market trading - e.g. London, NASDAQ and the NYSE – have a significant proportion of their trading through the limit order book. Electronic limit order book markets are much more

*sgupta24@fordham.edu, Gabelli School of Business, Fordham University

†v.raman@lancaster.ac.uk, University of Lancaster

‡pyadav@ou.edu. Price College of Business, Oklahoma University

§Paper available from authors on request. The authors gratefully acknowledge the NYU Stern-NSE initiative on the Indian Capital Market for funding the project.

transparent than the dealer or non-electronic-order-matching markets they have replaced, and provide the opportunity to market participants to observe significantly greater information through the trading process than had hitherto been possible, and one important reason is that they are now able to observe not only consummated trades but also the distribution of market or limit orders that led to these trades. This paper develops an empirical measure of informativeness based on the preferences revealed through the market and limit orders posted by different categories of traders (hereafter the IPRO measure), analyzes and illustrates the richness and effectiveness of the IPRO measure relative to the simple linear measures that have been proposed earlier in the literature [Anand et al(2005), Kaniel and Liu (2006)], and finally, utilizes the IPRO measure to empirically investigate several interesting issues and test the associated hypotheses.

It is important to be able to estimate and analyze the relative informativeness of different orders and traders as manifested in the distribution of market and limit orders since information is the primary driver of asset values, and is incorporated into prices through the trading activity of informed traders. In this context, it is also important to be able to estimate the cross-sectional variation in informativeness across different traders, i.e. information asymmetry; and hence it is not surprising that there is an extensive literature on information asymmetry.

In an efficient market, the absolute value of the change in price of an asset (or the volatility of returns for a zero expected return asset) reflects the net effect of new information flowing into the market from all traders. But this does not help us infer the informativeness of a particular set of orders and traders. The probability of informed trading (PIN) for an asset can be estimated across all traders in the market under certain restrictive assumptions as per the framework developed by Easley and O'Hara (1997) and others. However, once again, the methodology cannot be utilized for a particular trader or a set of orders or traders, since liquidity suppliers have to be assumed to make zero profits trading with that group, and that is particularly untenable for any specific group in an electronic limit order book where liquidity suppliers typically do not know the identity of their counterparties before they trade. However, the positioning spread measure, as defined for example in Hansch et al. (1999), can potentially be used to estimate the informativeness of trades relative to the benchmark collective informativeness of liquidity suppliers, even in the context of limit order book markets; and a variant of the positioning spread measure can be used to determine the informativeness of orders or traders: e.g., Anand, Chakravarti and Martell (2000) (hereafter ACM) used the average change in the mid-point of bid and offer prices from the time of submission of an order to a point of time 30 minutes later to proxy for the information content in the order. Similarly, Ron and Kaniel (2006) (hereafter RK) use a non-parametric variant – whether the change in price following order submission matches the direction of the trade – to estimate the informativeness of orders of a group of orders and traders.

The primary aim of this paper is to non-parametrically utilize the significantly greater information available in the preferences revealed through the entire distribution of market and limit orders, including the information in the tails of the distribution, to develop and estimate our IPRO measure of the informativeness of a trader or a group of orders, rather than just utilize the information contained in the parametric or non-parametric average change in price conditional on the order (as in ACM and RK); and thereby be able to provide a much richer characterization of the such informativeness. In subsequent versions of this paper, we plan to compare the results of utilizing our IPRO measure with the results of using the simple non-parametric RK and parametric ACM averages. We shall do this mainly by analyzing the manner in which the OPRI, ACM and RK measures predict the cross-section of actual economic profits made by different traders. Finally, after establishing the nuanced effectiveness of our IPRO measure, we shall use it to investigate several important issues and hypotheses of widespread interest to academics, regulators and market participants.

The basic assumption behind our informativeness measure is simple: a trader must have acted in his best interest based on his information at the time of submitting his order and acted accordingly. For example, if a trader has submitted a limit buy order then the expected profitability from submitting such order based on his information must dominate that of submitting any other type(s) of order(s) like limit sell, market buy or market sell. Using this revealed preference of order choice at the time of submitting one's order, we aim to recover the informativeness of the trader based on how does his ex-ante information compare to that of the ex-post market value of the stock. If the trader is more informed then his ex-ante value and the ex-post value of the stock must be tied 'more' to each other in the joint distribution, relative to someone who is relatively less informed. We use the copula techniques as used in the statistics literature to characterize the informativeness based on the properties of the Archimedean family of copula. Intuitively, based on our informativeness measure, if a more informed trader has a higher ex-ante value, it is more likely that the ex-post value of the stock will be higher. A detailed description of the methodology and the characterization is given in the following sections and in the appendix.

In the context of the growing literature on the issue, we first empirically analyze the types of traders and the characteristics of orders that have higher informativeness, and particularly the relative informativeness of market and limit orders conditional on the type of trader (e.g. individuals vs. domestic or foreign financial or domestic institutional investors), the trading style (algorithmic traders, liquidity suppliers, momentum traders, day traders etc.), the order aggressiveness, and market conditions. In subsequent version of the paper, we plan to examine whether the competitive advantage of informed traders comes from timing or selectivity. We plan to test how the level of informativeness of traders present in the market at a particular time drives price changes.

We make important contributions to the literature on informativeness and information asymmetry. Most importantly, we develop a conceptual framework for defining and measuring the informativeness of a set of orders and traders that is significantly richer in at least two ways. First, it contextualizes the strategic placement of market or limit orders within the totality of other limit orders, so that we are examining not just the extent to which the agent's order differs from the future price of the asset (as in ACM or RK), but also the extent to which the agent could have placed less or more aggressive orders. Second, it uses a measurement framework that compares the informativeness of two agents not just in terms of first order stochastic dominance but also in terms of second order stochastic dominance, and more generally in terms of the precision of the information signal of these agents; where the precision reflects not just precision in estimating the variance of the information signal distribution, but also the precision in estimating the tails of the distribution. This is accomplished through a framework that formally ensures the monotone likelihood ratio property arguably necessary in an informativeness measure: i.e., when an information signal is higher, it makes a higher true value more likely. Specifically, we use a suitably formulated copula framework that captures the potentially non-linear dependence between the inferred information signal about the value of the asset, and the realized future value of the asset.

Next, and equally importantly, we note that even though asset pricing and market efficiency are heavily anchored in the saliency of information and informed traders, we have relatively little direct evidence about informed traders, about the nature of their competitive information advantage, and about how their informativeness and the overall information asymmetry in the market actually affects prices and trading. In this context, we first extend the growing literature on the information content of market and limit orders by empirically analyzing the types of traders and the characteristics of orders that have higher informativeness, and particularly the relative informativeness of market and limit orders conditional on the type of trader (e.g. individuals vs. domestic or foreign financial or domestic institutional investors), the trading style (liquidity suppliers, momentum traders, day traders etc.), the order aggressiveness, and the various salient variables that describe overall market conditions. Second, we examine whether the competitive advantage of informed traders comes from timing the market or selectivity in picking the right asset. Third, we test how the level of informativeness of traders present in the market at a particular time drives price changes. And finally, we investigate whether and how IPRO-based directly measured information asymmetry across traders drives volatility and trading volume.

Our empirical analyses are based on an extremely rich dataset from the National Stock Exchange (NSE) of India. This is a heavily traded electronic limit order book market, behind only NYSE and Nasdaq in terms of the number of trades per day. The dataset contains the totality of orders and trades, provides the coded identities of each trader, and indicates whether the trader is an individual, a domestic financial institution, a domestic corporate institution or a foreign institutional investor.

The rest of the paper is organized as follows. Section I develops our IPRO measure of informativeness proposed in the paper, and describes how its parameters can be estimated. Section 2 outlines the features of the dataset used in this paper. Section 3 describes our measure and provides a descriptive analysis of informativeness as estimated from our IPRO measure, and the two linear measures that have been used earlier in the literature – RK and ACP measures – and establishes the relative effectiveness and richness of the IPRO measure based on the actual profits earned by the cross-section of market traders in different assets. Sections 4 tests various hypotheses related to information and order choice. Section 5 uses a by-product of our measure can be used as a proxy for ex-ante heterogeneity of trader's beliefs and tests various hypotheses related to heterogeneity and momentum profits. Section 6 concludes. We plan to add sections on conditional informativeness of market and limit orders, whether the competitive advantage of informed traders comes from timing or selectivity, how the level of informativeness in the market drives price changes, and investigate how IPRO-based directly measured information asymmetry drives volatility and trading volume.

1.1 Our Informativeness Measure: IPRO

In this section we describe the general measure of informativeness that could be estimated from the order submission behaviors. If a trader has more information than what is revealed in the current traded price he would take positions accordingly. Thus if he believes that price is going to go up in a particular interval of time then he would submit a buy order and if he believes that prices are going to go down then he would submit a sell order. His notional profit from his information in a particular interval (say a day) relative to the current midquote may thus be measured as the difference between the mid quote after 1 day relative to the current mid quote. The principle behind the informativeness measure is simple: higher informativeness should lead to higher notional profit. Now it is possible that two traders may have similar information hence for both of them this notional profit would be positive. However none of them have perfect information about this notional profit and one of them has better information than the other though imperfect. Both of them would incorporate their information in his order submission. Every order has a particular probability of execution (one for market order and less than one for limit order). Let the expected notional profit be the notional profit weighted by this probability of execution.

Let the true value of a stock be V, which follows a continuous distribution $F(V)$. Let each trader i receives a signal v_i about the true value of the object from a distribution $F(v_i)$. Loosely speaking his information set not only consists of his private knowledge of this signal v_i but also his perception of who else could be trading and what kind of information his opponent traders have given his signal v_i (i.e.$F(v_{-i}|v_i)$. Traders are assumed to be risk neutral and submit orders accordingly given this information. His choice set consists of whether to submit, a limit buy order (LB), a market buy order (MB), a limit sell order (LS), or a market sell order (MS) or refrain from trading. Given his bid submission strategy he decides what should be the amount of limt buy (sell) price p_b (p_s) and limit buy (sell) quantity, q_b (q_s).

Note that the ex ante value of the trader is not observable. We only observe the choices made by the trader: the buy/sell choice, limit/market order choice and the limit price at which the order is submitted for a limit order. We use the observed choices to estimate the trader's value signal for a particular stock based on his order.

Let us introduce a few more notations to characterize the order choice of the trader. If v; the ex ante value of the stock as perceived by the trader P_m: current market price and P_b

be the limit buy price with a probability of execution of P_e.

A trader would like to maximize his expected profit while submitting his order. For example a trader who submits a limit buy order at price Pb, he wants to maximize his expected profit as $(v - P_b) * P_e$ per unit of the shares. The revealed preference argument dictates that a trader who submitted a limit buy order must be making higher expected profit than by submitting any other type of order (say limit sell, or market buy). This would give us the following sets of inequalities for different types of observed orders.

$(v - P_b) \times \Pr(execution) > v - P_m \Rightarrow$ submit a limit buy order instead of a market buy order

$(v - P_b) \times \Pr(execution) < v - P_m \Rightarrow$ submit a market buy order instead of a limit buy order

$(P_s - v) \times \Pr(execution) < P_m - v \Rightarrow$ submit a limit sell order instead of a market sell order

$(P_s - v) \times \Pr(execution) < P_m - v \Rightarrow$ submit a market sell order instead of a limit sell order

$$\Pr(\text{a limit buy order is placed by trader}) = \Pr(v < \frac{P_m - P_b \times P_e}{1 - P_e}) \tag{1}$$

$$\Pr(\text{a market buy order is placed by trader}) = \Pr(v > \frac{P_m - P_b \times P_e}{1 - P_e}) \tag{2}$$

$$\Pr(\text{a limit sell order is placed by trader}) = \Pr(v > \frac{P_m - P_s \times P_e}{1 - P_e}) \tag{3}$$

$$\Pr(\text{a market buy order is placed by trader}) = \Pr(v < \frac{P_m - P_s \times P_e}{1 - P_e}) \tag{4}$$

In case of the buy order, we see that the point of indifference is the variable, which takes the value :

$$NewVar = \frac{P_m - P_b \times P_e}{1 - P_e} \tag{5}$$

Similarly, for a sell order, the value is -:

$$NewVar = \frac{P_m - P_s \times P_e}{1 - P_e} \tag{6}$$

Note that we can estimate this NewVar based on the observed limit order book. We outline below the details of the procedure to compute the NewVar. For limit orders, we take Pm as the best bid price on the opposite side of the book, Pe as the predicted probabilities based on the observed order execution behavior as outlined below, and the limit price (Pb or Ps depending on buy or sell limit order). For market orders, we take Pm as the price at which the order got executed. If the order was executed through multiple trades, we took the mean price of the executed trades for that order. The Ps/Pb and Pe for the market order was taken as the best available limit order price at that point of time (on the same side of the book) and its predicted probability of execution.

A trader would said to have better information if his ex ante value of the stock v is related to the ex-post value of the stock in the monotone likelihood ratio sense: a trader is better informed if his ex ante signal is higher than other trader when the true ex post value is high. Intuitively, if a trader is better informed then his ex-ante valuation should be more "coupled" with the ex-post valuation of the stock (as measured by the stock price one hour or one day post the submission of the limit order). That is the joint distribution of the ex-ante value of the informed trader and that of the ex-post value will be "more coupled". This is measured by the copula parameter between the ex ante and ex post value in the joint distribution as described below.

Let us use the revealed preferred order by a trader who submits a limit sell order for illustrative purpose below to express the joint distribution of ex-ante and ex-post return based on the trader's values. The joint distributions of returns for the rest of the orders will follow accordingly. For a limit sell order, we can write

$$
\begin{aligned}
\text{Pr(a limit sell order is placed by trader)} &= \Pr(v > \frac{P_m - P_s \times P_e}{1 - P_e}) \\
&= \Pr\blacksquare[v > NewVar)] \\
&= \Pr[-v \leq -NewVar] \\
&= \Pr[-v - P_L \leq NewVar - P_L] \\
&= \Pr[-(v + P_L) \leq -(NewVar + P_L)] \\
&= \Pr[\frac{-(v + P_{L)}}{P_L} \leq -\frac{(NewVar + P_L)}{P_L}] \\
&= \Pr[R_{Exante} \leq -\frac{(NewVar + P_L)}{P_L}] \\
&= P_{ex-ante} \qquad (7)
\end{aligned}
$$

The $P_{ex-ante}$ as defined above is the ex-ante probability distribution of the return that a trader who submits a limit sell order expects to receive, where P_L is the mid-point of the stock at the time of submitting the order. We estimate the probability of the ex-post actual return of the stock as: $\Pr[R_{ex-post} \leq \frac{P_{ex-post} - P_L}{P_L}]$ based on the actual ex-post stock price $P._{ex-post}$. The association between the joint distribution of the returns ex ante and returns ex-post will be used to compute the measure of the informativeness of a trader. We shall use the copula parameter between the two returns as the measure of association which will in turn be used for computing the measure of informativeness as defined in the appendix. As evident based on the above derivation, the joint distribution of returns is equivalent to the joint distribution of the ex ante value and ex-post prices.

2 Data Variable Descriptions and Estimation

We use the limit order book obtained from the National Stock Exchange of India (NSE) to estimate our measure of informtiveness. The National Stock Exchange (NSE) was created in 1994 as part of major economic reforms in India. It operates as pure electronic limit order book market, and uses an automated screen based trading system called National Exchange for Automated Trading (NEAT), which enables traders from across India to trade anonymously with one another on a real-time basis using satellite communication technology. NSE was the first exchange in the world to use satellite communication technology for trading.

In terms of total number of trades, NSE is the second largest pure electronic LOB market in the world, just behind Shanghai Stock Exchange (SSE), and it is the fourth largest among all markets irrespective of market structure, behind NYSE, NASDAQ and SSE. NSE 's order books accommodate all the standard types of orders that exist internationally in order-driven markets, including limit orders, market orders, hidden orders, stop-loss orders, etc. Limit orders can be continuously cancelled or modified without any incremental fees. NSE operates a continuous trading session from 9:55 am until 3:30 pm local time. The tick size is INR 0.05 (less than USD 0.01). Outstanding orders are not carried over to the next day. There is no batch call auction at the beginning of the trading day. The opening price is also determined by pure order matching.

Our sample consist of all the 50 stocks in Standard & Poor's CNX Nifty index, which represents about 60% of the market capitalization on the NSE and covers 21 sectors of the economy. The sample period is from April 1 through June 30, 2006, covering 56 trading days. Table 1 presents summary statistics on the trading characteristics of the sample stocks over the sample period. The mean market capitalization of the 50 stock in the sample is $7 billion, indicating these are relatively large stocks. There are, on average, 1,303 trades every 30 minutes, or approximately 43 trades per stock per minute. There are, on average, 1,678 order submissions per stock every 30 minutes, or about 56 order submissions per stock per minute. Further, the Bid-Ask spread, estimated from the order book and expressed as a ratio of the mid-quote, is about 3 basis points on an average. In sum, the 50 stocks that make up our sample are relatively large and liquid stocks.

Insert Table1 about here

The dataset provides complete information of trades and orders that enables the reconstruction of the order book to obtain best quotes and depth information. Further, the data also provides identification codes and classifications of traders for all the orders and trades in the dataset. We aggregate the 14 trader classifications flagged in the dataset into 4 broad categories: Individuals, Financial Institutions, Dealers, and Other Institutions. Table 2 presents descriptions of the four trader categories. While Individuals outnumber other trader categories, institutional traders, especially Dealers, are more active in terms of order submissions. Although the NSE is a pure electronic limit order book market with no designated intermediaries, Dealers, who are registered members of the NSE, trade on behalf of their clients and also trade for their proprietary accounts. These traders generally function as voluntary intermediaries at the exchange. More importantly, the dataset also provides identification codes of traders for all the orders and trades in the dataset, thereby enabling us to accurately estimate trader profitability over time and across stocks.

Insert Table2 about here

3 Using Copulas to Estimate Average Informativeness

In this section we define our measure of informativeness based on the estimated copula parameter between the two joint distribution of returns as described above. We can use this measure for a particular category of trader (e.g. algorithmic, individual etc.) or a particular trader (if we know the identity of the trader and can follow the trader's order submission behavior). We can estimate this measure for different interval of time like 1 hour, one day or seven day. We shall use the realized prices are the mid-point of the bid-ask prices at the relevant time.

Our measure of informativeness is based on the monotone likelihood ratio property (MLRP) between two random variables. Intuitively, if one trader has higher informativeness then he should receive higher (lower) ex-ante value signal when the ex-post value is higher (lower) relative to a trader who has lower informativeness. We characterize this notion in terms of the conditional distribution of ex-post values in terms of the ex-ante values. Equivalently, conditional on a higher value drawn from a more informed trader's distribution will make the conditional probability distribution of the ex-post value (conditional on the ex ante value) more likely. This property is equivalent to the monotone likelihood ration property in statistics. Specifically, MLRP implies

$$f(s_1, y_1)f(s, y) - f(s, y_1)f(s_1, y) \geq 0$$

whenever

$$s_1 \geq s, y_1 \geq y$$

The joint distribution of the ex-ante and ex-port values (and returns as shown earlier) of a stock $(F(R_{ex-ante}, R_{ex-post}))$ can be fully characterized by the copula parameter

$\alpha; F(R_{ex-ante}, R_{ex-post}; \alpha)$, where α is the degree of association between the two random variables $R_{ex-ante}$ and $R_{ex-post}$ (equivalently between ex-ante value v and ex-post value P_L as described earlier). A detailed description of various copula families and major properties are relegated to the appendix. It is shown in the appendix that their is a one-to-one mapping between the copula parameter and the average informativeness of as defined above. Specifically, for the Clayton family of copula average Informativeness is $\alpha_{Clayton}/(2*(2+\alpha_{Clayton}))$. Similarly, for Gumbel copula, average Informativeness is $(\alpha - 1)/(2 * \alpha_{Gumbel})$. Detailed derivations are relegated to the appendix.

We shall therefore estimate the copula parameter α between the ex-ante and ex-post return ($R_{ex-ante}$ and $R_{ex-post}$) for each trader (or for each trader type: algorithmic, institutional etc.) for a given horizon. We shall then compute its average informativeness and use for various comparisons as described below. We shall first estimate the copula parameters for various copula types (Frank, Clayton, Gumbel) for each trader and choose the best fitted copula based on a model selection criterion. We shall then estimate the average informativeness based on the best fitted copula parameter.

3.1 Estimation of the Copula Parameter

To estimate the copula parameter for each trader and every order, we need the order price (limit buy(P_b) or sell (P_b) or market depending on the order type), the prevalent market price (P_m) and the probability of execution of each order (P_e). We take the observed value of P_b, P_s and P_m for each order and estimate the ex-ante probability of execution of each order (P_e) based on the existing market condition at the time of submission of the order.

For a particular order, the ex post probability of execution can be computed as follows -:

$$Probability\ of\ execution(P_e) = \frac{Volume\ of\ order\ executed}{Volume\ of\ order\ placed}$$

Based on this, we can see that an order may either be fully executed ($P_1 = 1$), partially executed ($0 < P_e < 1$) or remain unexecuted ($P_e = 0$). Market orders always have $P_e = 1$. We can always compute the ex-post probability of execution of any order based on the actual volume executed. We use the logistic distribution to estimate the ex-ante probability of execution. We model partial and full executions separately.

We use the following set of independent variables to control for market condition to estimate the ex-ante probability of execution via a logistic regressions: order price aggressiveness, where limit buy order aggressiveness is defined as (limit buy price – best sell price) / best sell price and limit sell order aggressiveness defined as (limit sell price – best buy price)/best buy price. Market orders are considered most aggressive as they are always executed at the best price available on the opposite side of the book. We calculated aggressiveness of an order based on the five best prices on the opposite side of the book. This was done in order to take care of the depth at the best prices. For this reason, we also took into consideration the volume available at each of the five best prices. We did this for all the fifty stocks in our sample for both limit buy orders and limit sell orders. We also use depth ratio defined as the five of the most aggressive orders are considered while calculating both sell-side and buy-side depths. Specifically, buy order depth ratio is defined as buy-side depth /sell-side depth and sell order depth ratio is defined as sell-side depth /buy-side depth. Further, we also include the bid-ask spread, log of the order volume, trader category dummies, time of the day dummies to account for the unusual trading activity in the first and last thirty minutes of daily trading (open and close), and day fixed effects. The models are run separately for each of the 50 stocks in the sample. Aggregated results are presented in Table 3. The estimated coefficients are quite similar for partial and full order executions. Also, all the independent variables included are statistically significant. We find that the probability of order execution significantly increases with order aggressiveness; and reduces with bid-ask spread, depth ratio, and order volume. Further, probability of execution is significantly lower if the order is placed either in the first or the last 30 minutes of trading. Finally, we also find that order placed by exchange members and financial institutions are most likely to be executed, while orders placed by individuals or other institutions are less likely to be executed. We use these estimated models to compute predicted probability of execution for each order placed for all the fifty stocks in our dataset.

Insert Table3 about here

3.2 Trading Profitability and Informativeness Across Trader Categories

In this section, we exploit one of the unique features of our dataset to examine how our measure of informativeness (IPRO) and trader profitability varies across different trader categories discussed previously. The trader profitability is defined in the following way: Let $RI_{j,t}^{J}$ denote the level of rupee-inventory of trader j in stock i and time t, and let $RT_{j,t}^{J}$ denote the profit from round traders by trader j in stock i between time $t-1$ and t. We calculate trader profitability ($Total\_PL$) as follows:

$$Total\_PL_{i,t}^{J} = RI_{j,t}^{J} + RT_{j,t}^{J}$$

Further, to account for the vast differences in capital employed by different traders, we normalize trader profits by the average size of rupee inventory maintained by the trader. This normalized measure is denoted by $Total\_PL\_ratio_{i,t}^{J}$.

$$Total\_PL\_ratio_{i,t}^{J} = \frac{Total\_PL_{i,t}^{J}}{\overline{RI_{i}^{J}}}$$

Table 4 presents descriptive statistics of Total_PL, Total_PL_ratio, and our measure IPRO based on the Clayton copula. All three are estimated for each trader-stock combination, using data sampled at 30-minute intervals; their volume-weighted averages (across trader-stock combinations) are reported in Table 4. As seen from the table, the median trader losses money – median Total_PL and Total_PL_Ratio are both negative. Clearly, there is significant variation in trader profitability. Similarly, the copula based informativeness also shows great variation. Informativeness of all the trader-stock combinations are plotted in Figure 1. Informativeness ranges between -0.12 and 0.19. Moreover, similar to trader profitability, the mean and median informativeness are both negative.

Insert Table4 about here

Table 4, Panel A describes the relationship of trader ex-post profitability with ex-ante informativeness. Based on mean and median IPRO numbers, exchange members appear to be the most informed, followed by financial institutions, other institutions, and, finally, Individuals. Also, exchange Member IPRO is significantly greater than zero, but financial and other institutions are statistically indifferent from zero. However, individuals are clearly less informed – IPRO is significantly less than zero. Moreover, Exchange Members are also the most positively skewed of all trader categories, followed by financial institutions, individuals and other institutions. Interestingly, comparisons of trader profitability (Total_PL and Total_PL_Ratio) also yield similar results. Again, exchange members are the most profitable followed by financial institutions. In contrast to IPRO results, we find that Individuals are more profitable than other institutions.

IPRO distributions of different trader categories are presented in Figure 2, and a formal comparison of IPRO distributions of all trader categories are presented in Table 4, Panel B. We employ the nonparametric Dwass, Steel, Critchlow-Fligner Method (DSCF) procedure to compare different trader-category pairs of IPRO distributions. As expected, a comparison of distributions yields different results from a comparison of means. One, we find that IPRO distribution of Individuals continues to be significantly different from those of Exchange Members and Financial Institutions, but the differences in distribution of Individuals and Other Institutions are marginally insignificant. On the other hand, Other Institutions, Financial Institutions and Exchange Members have statistically similar IPRO distributions – p-values associated with their pairwise comparisons are all greater than 0.10.

3.3 Portfolio Sorts

In order to validate our measure of informativeness, we examine how it relates to trader profitability. First, we form portfolios based on deciles of Total_PL_Ratio and analyze how informativeness varies across these portfolios. The results of this analysis are presented in Table 5; Panel A reports in sample numbers of Total_PL, Total_PL_Ratio and IPRO, and Panel B reports out of sample values of the same variables.

Results in Table 5. Panel A and B both show a clear positive relation between trader profitability and the copula based measure of informativeness - greater trader profitability is associated with higher measures of informativeness. For example, when the average profitability (Total_PL_Ratio) is -1.16%, average informativeness is -0.0027; and when average profitability is 0.82%, average informativeness is 0.0007. Further, the relation between the two is almost monotonic. Also, the relation is similarly strong when measured through mean or median values.

3.4 Quantile Regressions

Next, we examine the same relation using Quantile regressions. Such an analysis is useful in understanding how the relationship varies across the distribution of trader profitability. The following regression is estimated at difference quantiles of trader profitability (Total_PL_ratio).

$$Total\_PL\_ratio_{i,t}^{J} = \alpha + \beta IPRO_i^t + \epsilon$$

Figure 3 plots the $\alpha's$ and $\beta's$ for each of the quantile examined. The figure depicts two important findings. One, the relation between trader profitability and informativeness is positive and significant across all the considered quantiles. Two, the relation between the two is especially strong in the tails. In other words, unlike most linear measures of informativeness, the copula based measured is most useful in explaining extreme values of trader profitability.

Insert Figure 3 about here

4 Informativeness and Order Choice

Whether an informed trader uses limit order or market order is an important policy question. Prior academic literature assumed that informed traders use market order to take advantage of their informativeness. However Bloomfeld, Ohara and Saar(2006))in their experimental paper documented that informed traders mostly use limit order. To test the hypothesis about the order choice we run a logistic regression on the observed order choice of a trader on their informativeness and other controls. The dependent variable is a binary variable equal to 1 if order aggressiveness is greater that 0.5 for a given trader in a given stock over a thirty minute interval. Results are presented in table 6.

We find a positive a significant coefficient of the informativeness variables. Thus it is more likely that a trader submits market order if he is more informed. The effect is also economically significant; one standard deviation of increase in informativeness makes it 2.8% more likely to submit a market order. However we find that the relationship between order choice and informativeness is more nuanced depending on the type of the information event or time of the day. For example, in the first hour of the day the likelihood of submitting a market order by a more informed trader goes up by about 2.9%. However during the post earnings period the likelihood goes down by about 1.4%.

We also looked at the relationship between order choice with that of a major macro economic shock. On May 18th 2006, Indian stock market had the largest ever intra-day drop in their histories so far with Sensex losing 826 points. Analysts speculated that the reason for this drop was that the US CPI number which was released the day before, was much above expectations. This couple with the weakness observed in the London Metal Exchange led to losses in emerging markets like India, Mexico and Brazil. We found that in the event of such a large macro shock informed trader become even more aggressive in their order choice. Specifically, for one standard deviation change in informativeness, the likelihood of submitting a market order goes up by about 2.2% after the macro event (on top of the 2.8% figure mentioned above).

Insert Tables 6 about here

We next study the role of informativeness on the choice of hidden orders. If a trader is more informed it is likely that he/she may prefer to hide his information behind the veil of hidden orders. We create a variable called *Hidden* ratio as the ratio between hidden limit order volume to the trader's total limit order volume. We run a logistic regression on informativeness where the the dependent variable takes value 1 if hidden ratio is greater than 0.5 or zero otherwise. Results are reported in table 7. The coefficient is significant and positive. The magnitude is also economically significant. A one standard deviation increase in informativeness increases the likelihood of submitting a hidden order by about 3.6%.

The choice of hidden order is also nuanced depending on the time of the day or type of information event. During the first hour of trading, for one standard deviation increase in informativeness, the likelihood of submitting a hidden order goes up by 6.6%. During the last hour of trading, for one standard deviation increase in informativeness, the likelihood of submitting a hidden order goes up by 3%. During a global macro shock, like the event on May 18th, for one standard deviation increase in informativeness, the likelihood of submitting a hidden order goes up by 3%. While during the post earnings announcements period, for one standard deviation increase in informativeness, the likelihood of submitting a hidden order goes down by 3.5%.

Insert Tables 7 about here

4.1 Informativeness and Trader Type: Algo vs. Non-Algo

The algorithmic trading was allowed in India since 2008. Our dataset for the period 2012 has a flag whether the trader is an algorithmic trader or not besides trader categories (institutional, exchange members etc.). In table 8, we compare the informativeness of various trader categories. Panel A and B describes the composition of various types (institutional, individual, exchange members) of traders into algorithmic and non-algo types and their respective participation and liquidity provision. In Table 9, we find that within every category (institutional, individual and exchange members) the algorithmic traders are more informed than non-algo traders.

Insert Tables 8 and 9 about here

5 Heterogeneity of Traders Values and Trading Styles

One of the by-product of our measure of informativeness is the bound on the ex-ante value perceived by the trader while submitting their order. The bounds on this value is represented in equation 5 and 6 for buy and sell orders respectively. This value is therefore available for every trader for every order. We use this as a proxy for the ex-ante belief about the value of the stock as perceived by the trader while submitting their order. Heterogenous beliefs and its impact on speculation and trading in the financial market is an important policy question (see the survey by Scheinkman and Xiong (2004)). Miller (1977) for example argued that if agents have heterogenous beliefs about an asset's fundamental and short sales are not allowed, equilibrium priced would reflect the opinion of the more optimistic investor.

Using the ex-ante bound on the value as a proxy for a trader's belief about his ex-ante value, we define two measures of heterogeneity of beliefs of traders at a particular interval during trading hours. We define coefficient of variation of the ex-ante values calculated every minute for all traders and then averaged over every thirty minutes of trades and then averaged over a day to get a daily measure of heterogeneity of beliefs. We similarly use interquintile range $\{(75th$ Quintile $-25th$ Quintile$)/75th$ Quintile$\}$ as another measure of heterogeneity of beliefs over the day. Their behavior is plotted in figure 5.n One interesting point to note from this graph is that the heterogeneity of beliefs goes up significantly around the global macro event period (in the middle region of the graph). In Panel A of figure 5, we plot the intra-day pattern of this heterogeneity. The graph has an interesting U-shaped pattern, signifying higher heterogeneity of beliefs toward the opening and close of the day.

Insert Figure 5 about here

In the following table (Table 10), we link the trader heterogeneity with momentum profits. Stocks are sorted based on their prior 30-minute returns into three momentum categories (High, Moderate and Low). We then report a double sort of this momentum profit along with trader heterogeneity (low, moderate and high heterogeneity) into portfolios in table 10. The low return and low heterogeneity portfolio sort has a value of 0.19, signifying that if returns are low then it turns positive 0.2 standard deviation of return.

Insert Table 10 about here

In table 11 we analyze the relationship between intraday momentum with the heterogeneity of beliefs. The table shows the Fama-Mcbeth regression of predictive cross-sectional regression for thirty minute stock returns. The dependent variable is the mid-quote based return of thirty minutes interval. It is regressed on the interquintile based measure of trader heterogeneity. The coefficient of the trader heterogeneity is positive and significant suggesting a positive momentum profit over a thirty minute interval as the trader heterogeneity increases. This is consistent with Verardo (2009) and Allen Morris and Shin (2006).

Insert Table 11 about here

6 Conclusion

In this paper, we have proposed a new measure of ex-ante informativeness of a trader based on his order submission behavior. We apply our measure to the limit order book from the National Stock Exchange (NSE) of India. We use the informativeness measure to predict the ex-post profitability of the trader. Our measure capture extreme values of trader profitability suggesting non-linearity of informativeness. Our measure performs better out of sample relative to other measures like Anand et al and Kaniel and Liu. We test various hypothesis related to information and order submission using our measure and provide quantitative estimate of the relationship. Finally, we use heterogeneity of beliefs as a by-product of our measure and test various hypothesis related to momentum profits and heterogeneity of beliefs.

7 Appendix

8 Modelling Informativeness: Copula

For expositional clarity and notational simplicity we shall describe the concept of copula for the case of two variables only. However, as will be clear later, the concept extends well for n variables.

As described in the definition, informativeness describes the 'interlinkage' between two random variables, i.e., if one bidder receives a high signal then (positive) informativeness means that the likelihood of the true price being high is high. Note that independence is a special concept of informativeness . Thus informativeness should measure how are the order S and the true value Y are related in the joint distribution $F(s, y)$. Let the continuous marginal distribution of order and the true value be F_S and F_Y respectively. Note that if the distribution of the limit price and the true value are same then F_S and F_Y are same distributionally otherwise they are different. Let $U = F_S(S)$ and $V = F_Y(Y)$. U and V are called 'positive integral transformation' of random variables S and Y respectively. Note that, U and Y are also random variables and let u and v be the values these random variables take[1].

Then according to Sklar's ('59) theorem there exists a unique copula function $C(u, v)$ that 'couples' or joins multivariate distribution functions to their one dimensional marginal distribution functions such that $F(s, y) = \Pr(S \leq s, Y \leq y) = \Pr(F_S(S) \leq F_S(s), F_Y(Y) \leq F_Y(y)) = \Pr(U \leq F_S(s), V \leq F_Y(y) = C(F_S(s), F_Y(y))$.

Thus $C(.,.) = F(.,.)$ is a distribution function that connects $F(s, y)$ to the marginals F_S and F_Y respectively. Thus for general case, copula is a map from $[0, 1]^k$ to $[0, 1]$, where k is the dimension of the joint distribution.

Definition 1 *A copula is the distribution function of a random variable* $S = (S_1, .., S_k)$ *in* $\Re^k$ *with uniform*$(0, 1)$ *marginals*[2].

More discussion and specific examples on copulas can be found in Joe ('97) and Nelson ('99). A simple example of copula is independent copula, defined as $\Pr(S \leq s, Y \leq y) =$

[1] Throughout this paper, we use upper case letters to denote the random variables and lower case letters to denote the value it takes.

[2] Heuristically, $\Pr(U \leq u) = \Pr(F_1(S_1) \leq u) = \Pr(S_1 \leq F_1^{-1}(u)) = F_1 F_1^{-1}(u) = u$. This is the distribution function of a uniform random variable.

$F_S(s) \times F_Y(y) = u.v = C(u,v)$. The following theorem by Sklar ('59) establishes the existence and uniqueness of copula representation.

Theorem 2 *(Sklar) Let S and V be random variables with distribution functions F_S and F_V respectively, and joint distribution function F. Then there exists a copula C such that $F(s,y) = C(F_S(s), F_Y(y))$. If F_S and F_Y are continuous then C is unique.*

Thus according to Sklar as long as we can identify the marginal distributions F_S and F_Y we can identify the joint distribution of order by the copula function.

Also copula is invariant under increasing transformations[3] :

Theorem 3 *If S and Y has copula C_{SY} and $W(s)$ and $V(y)$ are strictly increasing functions then $W(s)$ and $V(y)$ also has copula C_{XY}.*

Thus S and Y have the same copula function as U and V since U and V are positive integral transformations of S and Y respectively.

C can be parametrized by a parameter α, to have a specific functional form and be denoted by $C(u,v;\alpha)$ while keeping the marginals unspecified. Since $C(u,v;\alpha)$ is a distribution function, let the corresponding density function be $c(u,v;\alpha)$. Gaussian copula, Frank copula, Gumbel copula, Clayton copula as described in table 1 are examples of parametric families of copulas. The parameter α measures the degree of dependence (concordance) between the random variables u and v. We show below α as a measure of average informativeness between the order. Note that, by definition, informativeness is a dependence concept such that the joint density exhibits the total positivity property. A simple test of informativeness for continuously differentiable density functions is that the density function is log supermodular. Now consider the density function of say Frank copula given by,

$$c(u,v) = \frac{(\alpha-1)\log(\alpha)\alpha^{u+v}}{\{(\alpha-1)+(\alpha^u-1)(\alpha^v-1)\}^2}, \qquad 0 < \alpha < \infty, 0 < (u,v) < 1$$

It is easy to verify that $\frac{\partial^2 \log c}{\partial u \partial v} \geq 0$ for $\alpha < 1$, thus u and v are (positively) affiliated. For $\alpha > 1$, u and $-v$ (or $-u$ and v) are (positively) affiliated. Hence the Frank copula satisfies the informativeness property and hence the random variables u, v and hence x, y are affiliated. Note that, when $\alpha \to 1$, $c(u,v) \to 1$, i.e., then u and v are independent. Below

[3] A proof can be found in Nelsen('99).

we present a special group of copulas called Archimedean copulas. Archimedean copulas are those distribution functions $F(u,v)$, such that $F(u,v) = \tau^{-1}[\tau(u) + \tau(v)]$ for some convex, decreasing function τ.

We give below a few Archimedean copulas and their major properties

$Table 1$: Different Copula Families

Copula	$C(u,v)$	$\tau(v)$	Range of α	Total Positivity
Frank	$\log_\alpha[1 + \frac{(\alpha^u-1)(\alpha^v-1)}{(\alpha-1)}]$	$log_\alpha(\frac{1-\alpha}{1-\alpha^s})$	$[0,\infty)$	Yes
Clayton	$(u+v-1)^{\frac{1}{\alpha}}$	$\frac{v^{-\alpha}-1}{\alpha}$	$[0,\infty)$	Yes
Gumbel-Hugard	$\exp(-[(-\log u)^\alpha + (-\log v)^\alpha)]^{1/\alpha})$	$[-\log(v)]^\alpha$	$[1,\infty)$	Yes

The copula function thus can be parameterized by a copula parameter that couples the two random variables. Thus the joint distribution is identified by the copula parameter and the respective marginal distributions. Note that there is no restriction on the marginal distributions besides continuity. Thus if the marginal distribution can be estimated non-parametrically the estimation of the joint distribution is semi-parametric where the copula parameter is parametrically identified. The parametric assumption is on how are the two random variables related or coupled. The estimation procedure is thus semi-parametric.

There are obvious advantages of using the copula instead of the joint distribution. The joint distribution may be very complicated and specially hard to estimate using non-parametric procedures due to the 'curse of dimensionality'.[4] Instead while using copulas we only need to estimate the marginals which are unidimensional. The marginals can be nonparametric or parametric. Once the marginals are estimated, the estimation of the joint distribution boils down to specifying and copula functional form to estimate the copula parameter only using a maximum likelihood procedure. The entire estimation procedure is semi-parametric where the marginals are still estimated non-parametrically.

[4]Non-parametric estimation procedures like Kernels requires some kind of averaging over the data within a certain bandwidth. The computing cost of that averaging increases exponentially in the dimension of the random variable. This is refered in the literaure as 'curse of dimensionality'.

8.1 Copula as a Measure of Average Informativeness

By definition, for two variables, $s' \geq s, y' \geq y$, informativeness is equivalent to $f(s', y)f(s, y') \leq f(s, y)f(s', y')$. Standard manipulations yield,

$$\frac{f(y'|s)}{f(y|s)} \leq \frac{f(y'|s')}{f(s|s')}$$

Thus informativeness $\Rightarrow f(s', y')f(s, y) - f(s, y')f(s', y) \geq 0$, whenever $s' \geq s, y' \geq y$. Below we define average informativeness ,

$$T = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{y'}\int_{-\infty}^{s'} [f(s', y')f(s, y) - f(s, y')f(s', y)] ds dy ds' dy'$$

Let $u = F_1(x)$, $v = F_2(y)$,

$$T = \int_0^1\int_0^1\int_0^{v'}\int_0^{u'} [c(u', v')c(u, v) - c(u, v')c(u', v)] du dv du' dv'$$

where, $c(u, v) = \frac{\partial^2}{\partial u \partial v} C(u, v)$, $f(s, y) = c(F_S(s), FY(y))f_S(s)f_Y(y)$

Plug this back into $T(.)$, we get

$$\begin{aligned}
T &= \int_0^1\int_0^1\int_0^{v'}\int_0^{u'} [c(u', v')c(u, v) - c(u, v')c(u', v)] du dv du' dv' \\
&= \int_0^1\int_0^1 \{[\frac{\partial^2}{\partial u \partial v} C(u, v)]C(u', v')\} du' dv' - \int_0^1\int_0^1 \{\frac{\partial}{\partial u} C(u', v')\frac{\partial}{\partial v} C(u', v')\} du' dv' \\
&= \frac{1}{2}[4\int_0^1\int_0^1 C(u, v) dC(u, v) - 1]
\end{aligned}$$

For example, for Gumbel copula

$$T = \text{Average Informativeness} = \frac{\alpha - 1}{2\alpha}, \alpha \geq 1$$

Figure 1: IPRO across Trader Categories

This figure presents the distribution of trader informativeness across different trader categories (defined in Table 2). *IPRO* is the copula-based measure of informativeness estimated for each trader-stock combination using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during our sample period, April to June, 2006.

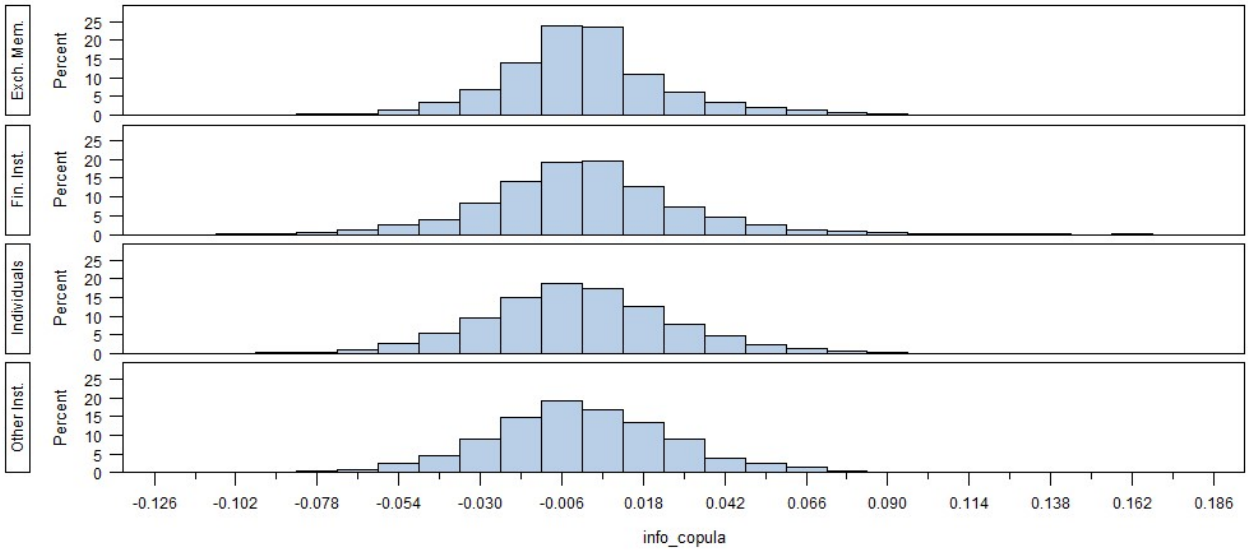
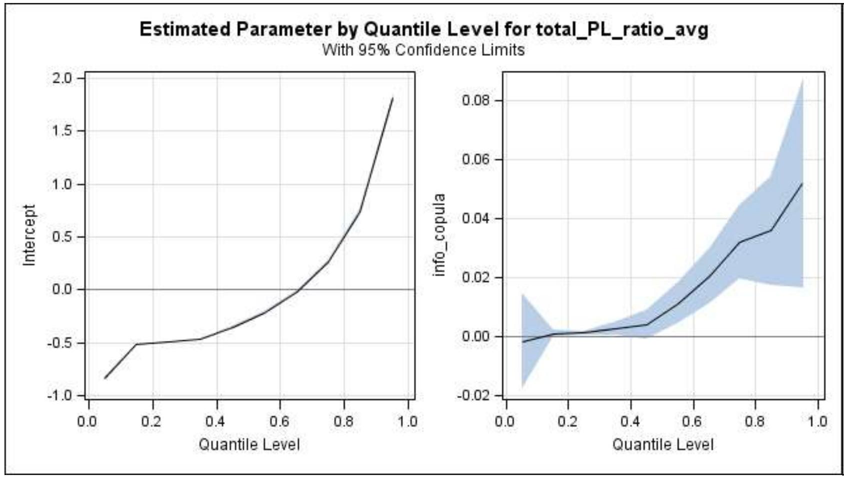


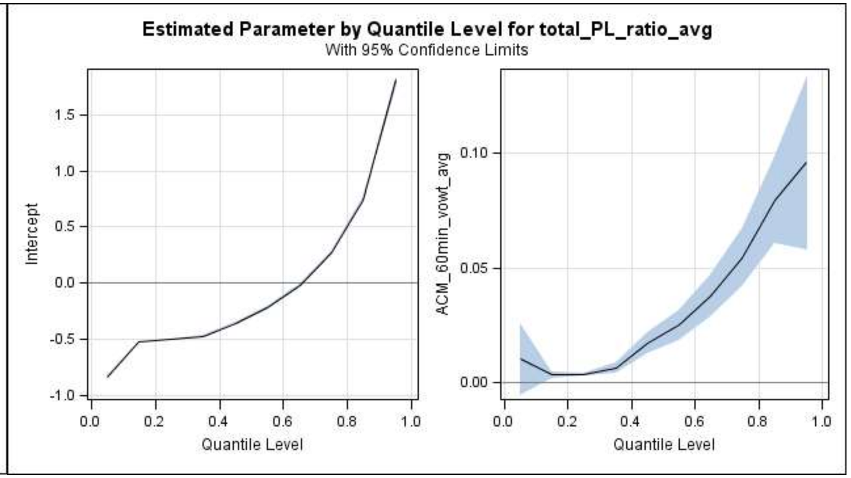
Figure 2: IPRO, Trader Profitability, and Trader Sharpe Ratios - In-Sample Quantile Regressions

This figure presents results from Quantile regressions used to examine the relation between different measures of trader informativeness and trader profitability. The analysis is conducted using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during the entire sample period, April to June, 2006. *Total_PL* is the average hourly profit/loss (sum of change in market value of inventory and profit/loss from round-trip transactions during the one-hour interval) estimated for each trader-stock combination. *Total_PL_Ratio* is the average of the ratio of hourly profit/loss and the capital) employed (absolute value of inventory by the trader in the stock during the one-hour interval. *Total_PL_Sharpe_Ratio* is the ratio of the difference between *Total_PL_Ratio* and the risk-free rate, expressed as a ratio of the standard deviation of *Total_PL_Ratio*. *ACM* is the average hourly measure of informativeness estimated (for each trader-stock combination) as proposed by Anand, Chakravarthi and Martell (2000). *IPRO* is the copula-based measure of informativeness calculated for each trader-stock combination.

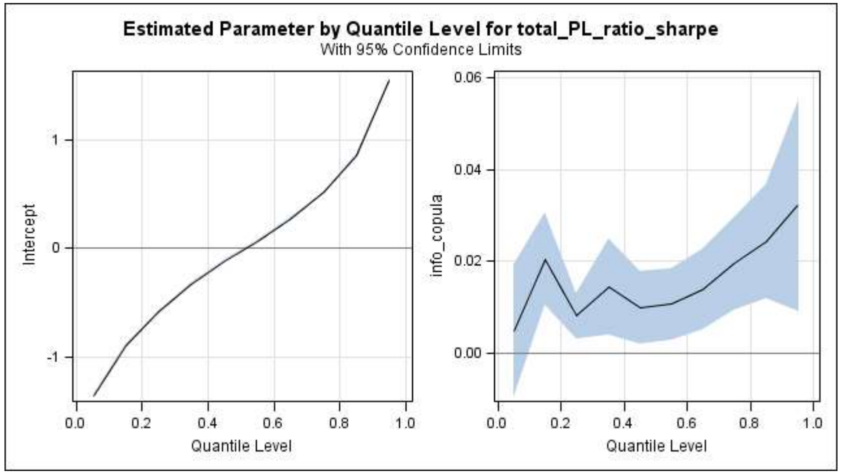
Panel A: IPRO and Total_PL_Ratio



Panel B: ACM and Total_PL_Ratio



Panel C: IPRO and Total_PL_Sharpe_Ratio



Panel D: ACM and Total_PL_Sharpe_Ratio

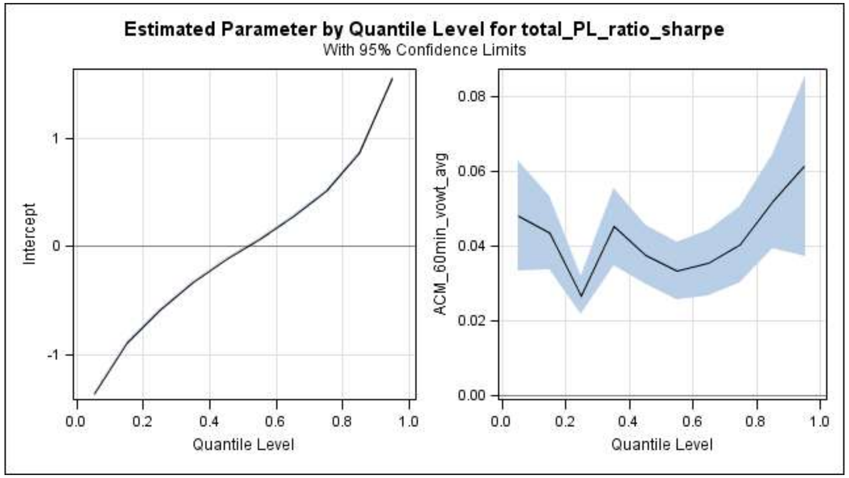
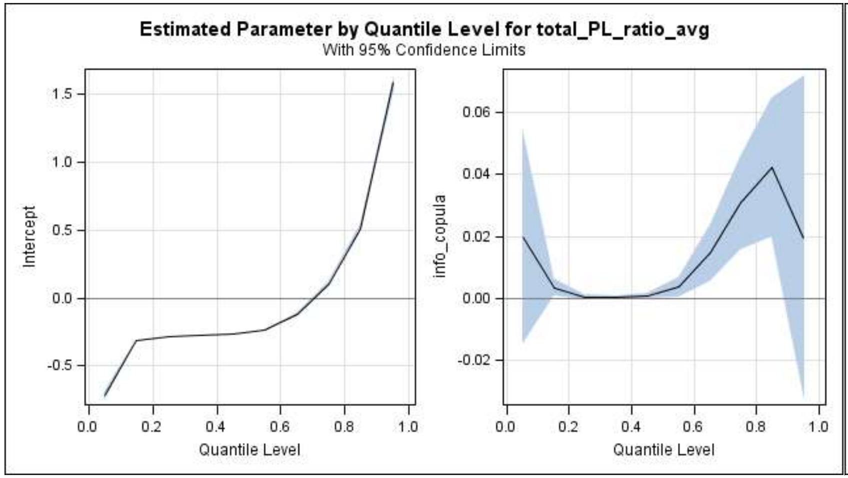


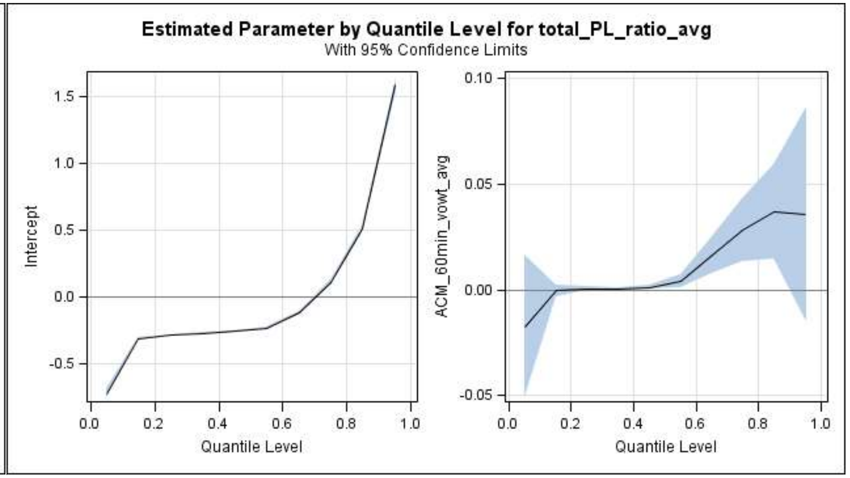
Figure 3: IPRO, Trader Profitability, and Trader Sharpe Ratios – Out-of-Sample Quantile Regressions

This figure presents results from out-of-sample Quantile regressions used to examine the relation between different measures of trader informativeness and trader profitability. The analysis is conducted using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during the entire sample period, April to June, 2006. *Total_PL* is the average hourly profit/loss (sum of change in market value of inventory and profit/loss from round-trip transactions during the one-hour interval) estimated for each trader-stock combination. *Total_PL_Ratio* is the average of the ratio of hourly profit/loss and the capital) employed (absolute value of inventory by the trader in the stock during the one-hour interval. *Total_PL_Sharpe_Ratio* is the ratio of the difference between *Total_PL_Ratio* and the risk-free rate, expressed as a ratio of the standard deviation of *Total_PL_Ratio*. *ACM* is the average hourly measure of informativeness estimated (for each trader-stock combination) as proposed by Anand, Chakravarthi and Martell (2000). *IPRO* is the copula-based measure of informativeness calculated for each trader-stock combination. *IPRO* and *ACM* are calculated using the first-half of the data (between April 1st and May 14th, 2006), and *Total_PL_Ratio*, and *Total_PL_Sharpe_Ratio* are calculated using the second-half of the data (between May 15th and June 30th, 2006).

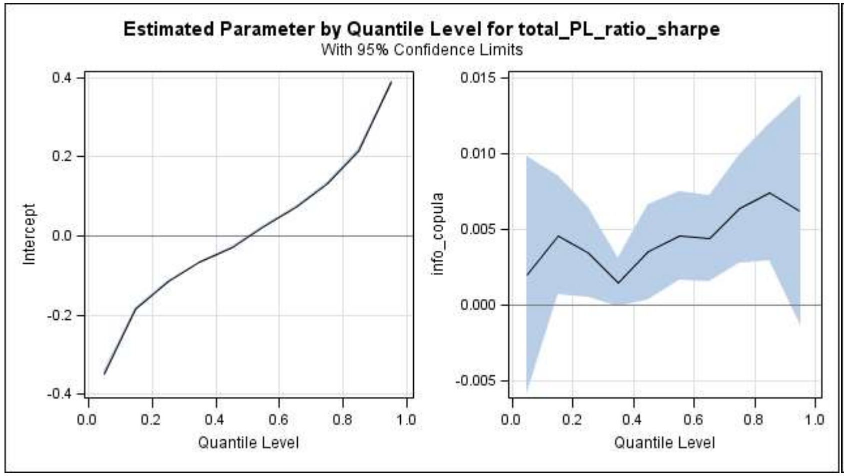
Panel A: IPRO and Total_PL_Ratio



Panel B: ACM and Total_PL_Ratio



Panel C: IPRO and Total_PL_Sharpe_Ratio



Panel D: ACM and Total_PL_Sharpe_Ratio

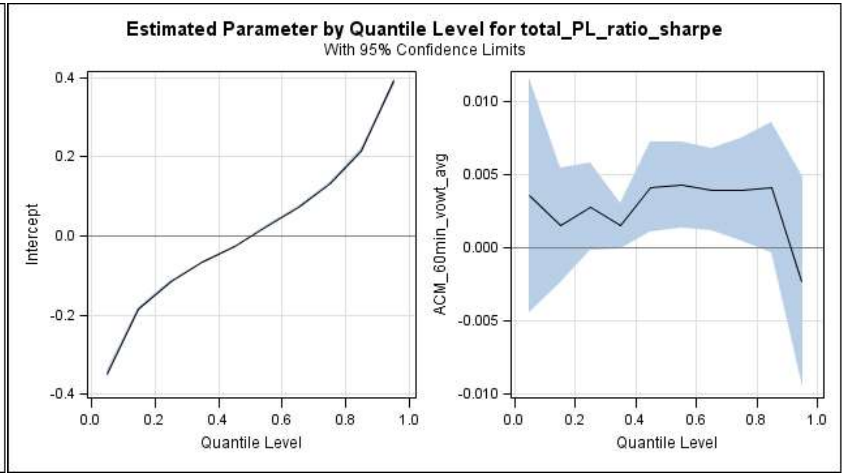


Figure 4: Informed Traders and the Forex Event

This figure compares the cumulative trading volume of informed and uninformed traders around the Forex event. The analysis is conducted using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during the entire sample period, April to June, 2006. *IPRO* is the copula-based measure of informativeness calculated for each trader-stock combination. Traders in the top quartile of IPRO are identified as informed and those in the bottom quartile are identified as the uninformed.

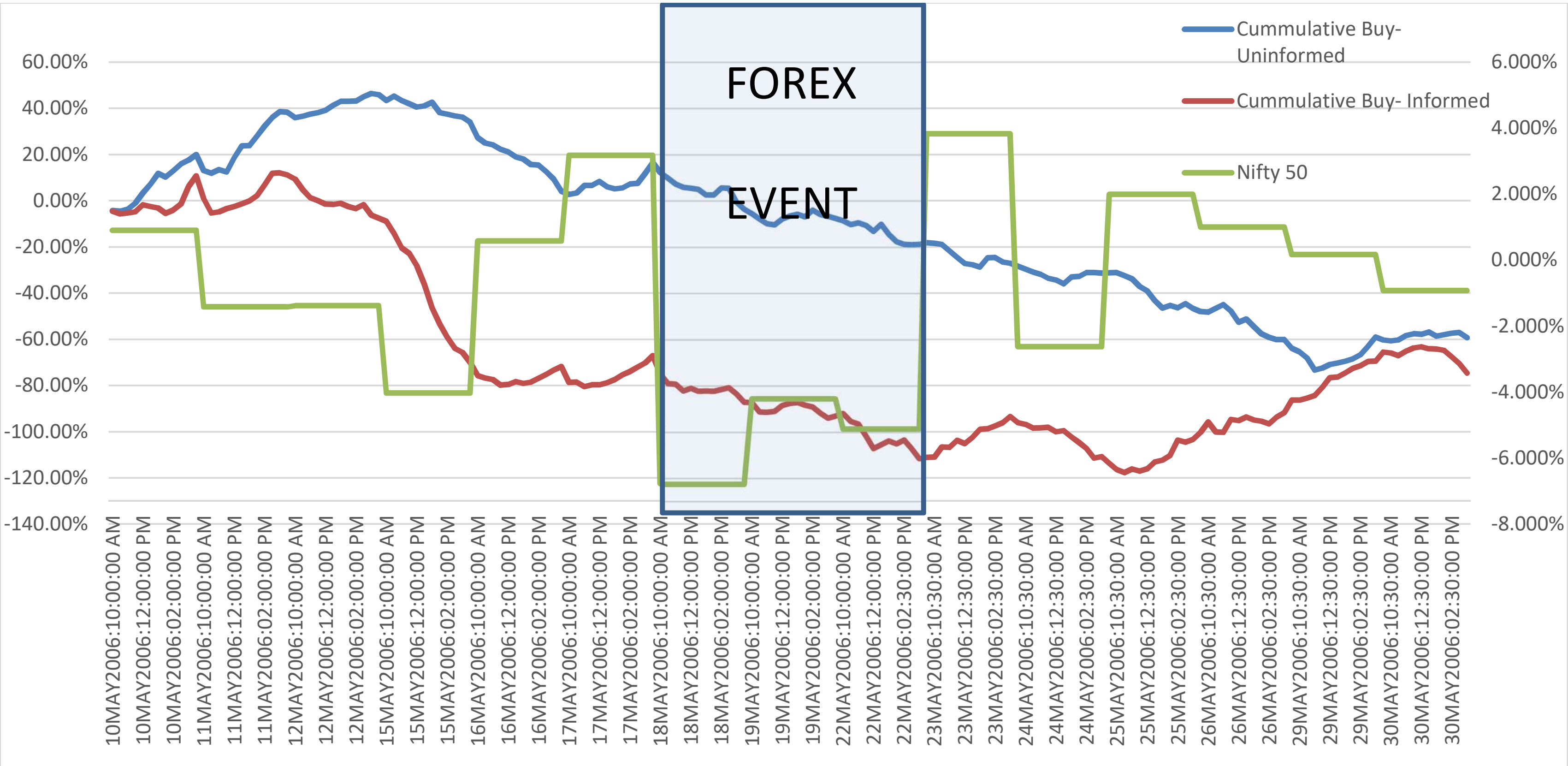
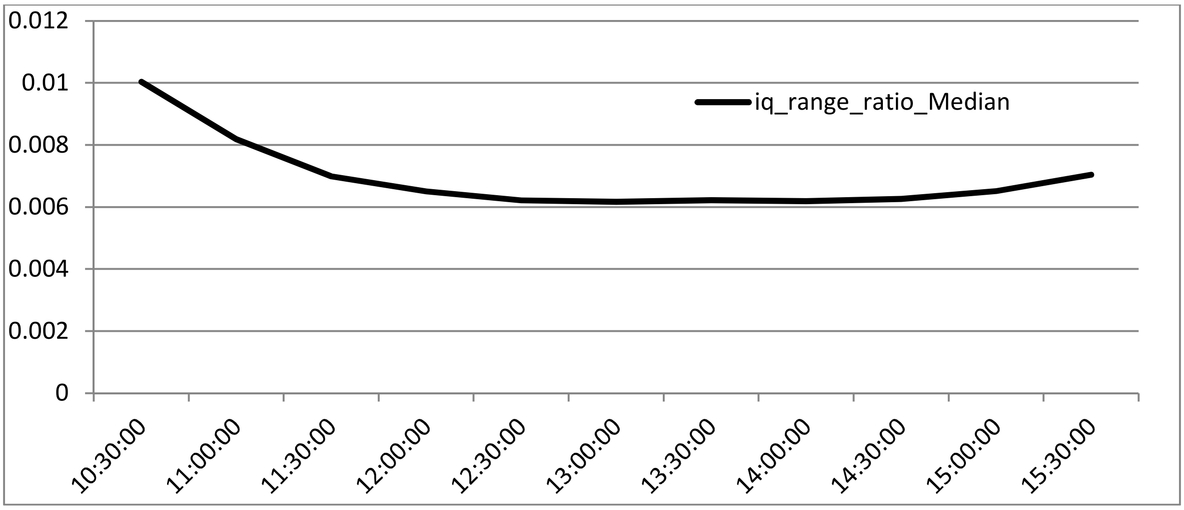


Figure 5: Variation in Heterogeneity

This presents variation in trader heterogeneity over a day and over the entire sample period. The analysis is conducted using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during our sample period, April to June, 2006. Traders' expectations of stock prices are extracted from their limit order sub missions as explained in section XXX. *IQ_Range_Ratio* is measured as the ratio of the inter-quartile range of traders' expectations of stock prices and the average of the expectations of stock prices. In Panel A, thirty-minute median values of *IQ_Range_Ratio*, calculated across the fifty stocks and over the entire sample period, are plotted. In Panel B, daily cross-sectional median value of *IQ_Range_Ratio* are plotted.

Panel A: Intraday Variation



Panel B: Variation over the Sample Period

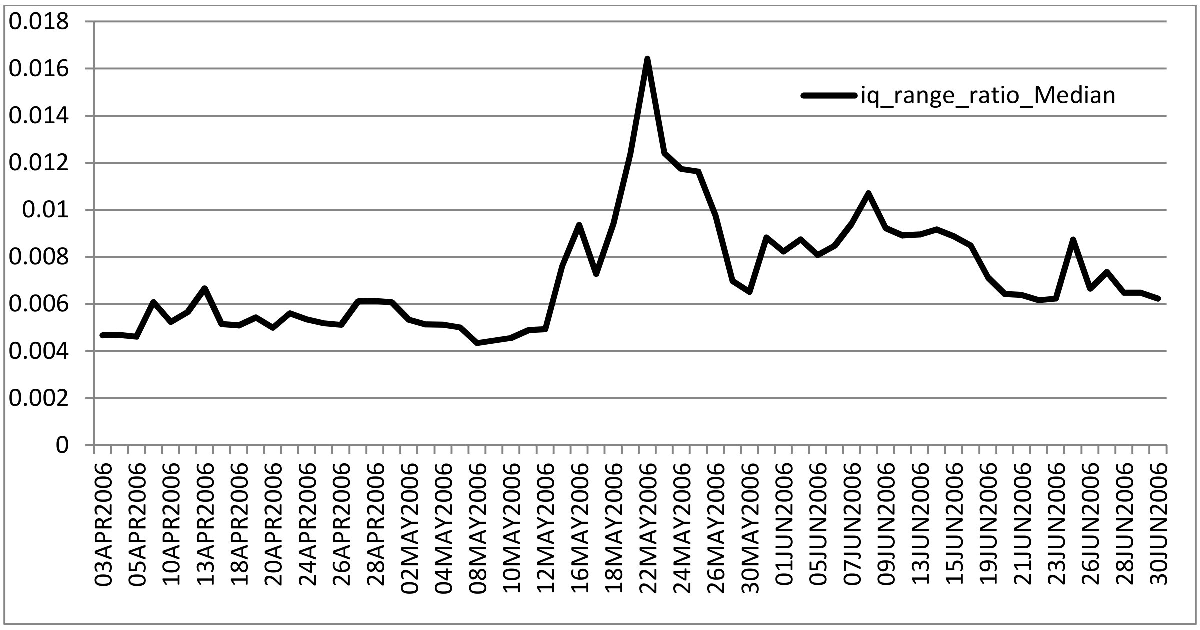


Table 1: Sample Descriptive Statistics

This table presents descriptive statistics of characteristics of 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during our sample period, April to June, 2006. *Panel A* presents descriptive statistics of stock characteristics. *Number of Trades* is the average number of trades in a stock in the sample; it is first calculated over 30 minute intervals for each stock and then averaged across the 50 stocks in the sample. *Volume of Trades*, *Number of Orders* and *Volume of Orders* are calculated analogously. *Buy Depth* and *Sell Depth* are the total volume of the ten most aggressive limit orders on the buy side and sell side of the book respectively. *BidAsk Spread* (estimated from the order book, expressed as a ratio of the mid-quote), *Return* (total stock return) and *Volatility* (standard deviation of *Return*) are first calculated over 30 minute intervals for each stock and then averaged across the 50 stocks in the sample.

	Mean	Median	Std. Dev.
Market Capitalization (USD Billions)	7	4	3
Number of Trades	1303	910	1165
Volume of Trades	121343	48294	174902
Number of Orders	1678	1150	1450
Volume of Orders	469357	207827	608518
Bid-Ask Spread	0.03%	0.02%	0.03%
Volatility	0.43%	0.42%	0.07%
Return	-0.02%	-0.02%	0.02%

Table 2 –Trader Categories

This table describes the different trader categories identified in the data. Their share of total limit order volume submitted in the sample, and the proportions of their limit order volume that are cancelled, modified and revised (cancelled or modified) are also presented. The proprietary data from the NSE identifies 14 different trader clienteles, which are further classified into 4 broader categories: Individuals, Financial Institutions, Dealers and Other Institutions.

Trader Category	Description	Number of Traders	Percentage of Total Limit Order Volume Submitted
Individuals	Individual Non-Residential Indians HUF (Families)	1,070,125	32.18%
Financial Institutions	Mutual Fund Bank Insurance Other Domestic Financial Institutions Foreign Financial Institutions	5,771	16.45%
Exchange Members	Dealers	509	40.68%
Others Institutions	Public and Private companies Partnership Firms Trusts and Societies Other Corporate Bodies Statutory Bodies	153,894	10.69%

Table 3: Probability of Execution

This table presents results from Logit regressions used to model the probability of execution of orders placed in the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India – during our sample period, April to June, 2006. The regressions are run for each of the 50 stocks separately; aggregated results and t-stats are presented. Regression results for *Partial Executions* (where the dependent variable is a binary variable equal to one if the order is partially executed) and *Full Executions* (where the dependent variable is a binary variable equal to one if the order is fully executed) are presented separately. For a buy order, *Aggressiveness* is defined as the percentage difference between the limit price and the prevailing best bid price; it is defined analogously for a sell order. *Bid-Ask Spread* is the difference between the difference between the best sell price and the best buy price; it is expressed as a ratio of the mid-quote. For a buy order, *Depth Ratio* is defined as the ratio of the buy-side depth (10 most aggressive orders) and the sell-side depth (10 most aggressive orders); it is defined analogously for a sell order. *Log_Volume* is the log of the order volume. *Open* is a binary variable equal to 1 for the first hour of daily trading. *Close* is a binary variable equal to 1 for the last hour of daily trading. *Fin Institutions* is a binary variable equal 1 if the order was placed by a financial institutional trader. *Other Institutions* and *Exch. Members* are also binary variables that are similarly defined. All regressions are run with day fixed effects.

	Partial Executions			Full Executions		
Variable	*Median*	*Mean*	*t-value*	*Median*	*Mean*	*t-value*
Intercept	2.07	1.98	28.02	2.19	2.12	28.40
Aggressiveness	72.12	76.95	34.48	72.37	76.97	32.33
Bid-Ask Spread	-47.58	-46.65	-23.83	-49.00	-49.99	-24.98
Close	-0.01	-0.02	-4.78	-0.03	-0.03	-8.73
Depth_Ratio	-0.03	-0.04	-10.80	-0.04	-0.05	-11.42
Log_Volume	-0.12	-0.12	-17.91	-0.24	-0.23	-33.05
Open	-0.02	-0.02	-2.91	-0.02	-0.02	-3.18
Fin. Institutions	0.15	0.13	11.68	0.21	0.20	15.02
Other Institutions	-0.13	-0.15	-8.07	-0.15	-0.19	-8.13
Exch. Members	0.27	0.25	16.04	0.42	0.40	25.04
Day Fixed Effects			Yes			

Table 4: Trader Profitability and Informativeness - Descriptive Statistics:

This table presents descriptive statistics of trader profitability, trader profitability ratio, and trader informativeness estimated using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during our sample period, April to June, 2006. *Total_PL* is the average hourly profit/loss (sum of change in market value of inventory and profit/loss from round-trip transactions during the one-hour interval) estimated for each trader-stock combination. *Total_PL_Ratio* is the average of the ratio of hourly profit/loss and the capital) employed (absolute value of inventory by the trader in the stock during the one-hour interval. *IPRO* is the copula-based measure of informativeness calculated for each trader-stock combination. Panel A presents overall descriptive statistics for the three measures over the entire sample; Panel B presents the descriptive statistics for the three measures across different trader categories defined in Table 2; and Panel C presents pairwise two-sided comparison of informativeness across different trader categories, along with Wilcoxon and DSCF test values.

Panel A: Overall

	Total_PL	*Total_PL_Ratio*	*IPRO*
Minimum	-42,381,387	-13.03%	-0.121
P5	-73,144	-0.77%	-0.046
Mean	81	-0.07%	-0.001
Std. Deviation	3,281	0.00%	0.000
Median	-588	-0.02%	-0.002
P95	72,291	0.49%	0.048
Maximum	75,614,331	12.56%	0.185

Panel B: Trader Profitability and Informativeness across Trader Categories

	Total_PL				*Total_PL_Ratio*				*IPRO*			
Trader Category	Median	Skewness	Mean	t-value	Median	Skewness	Mean	t-value	Median	Skewness	Mean	t-value
Exch. Mem.	2403.25	0.08	30994.13	0.84	0.01	1.48	0.02	1.32	-0.02	0.62	0.12	2.97
Fin. Inst.	-10.55	21.54	31352.41	0.87	0.00	1.87	0.02	0.89	-0.01	0.45	0.08	1.50
Individuals	-655.35	1.12	-3764.49	-5.45	-0.03	-2.65	-0.08	-27.93	-0.18	0.29	-0.09	-6.94
Other Inst.	-700.75	-31.37	-11156.01	-3.93	-0.05	-3.35	-0.14	-11.43	-0.10	0.19	0.00	-0.09

Panel C: Pairwise Two-Sided Multiple Comparison Analysis of Informativeness across Trader Categories

Trader Category Pairs	Wilcoxon Z	DSCF Value	Pr > DSCF
Exch. Mem. vs. Fin. Inst.	0.31	0.44	0.99
Exch. Mem. vs. Individuals	4.29	6.07	0.00
Exch. Mem. vs. Other Inst.	1.27	1.79	0.58
Fin. Inst. vs. Individuals	2.89	4.09	0.02
Fin. Inst. vs. Other Inst.	0.73	1.03	0.89
Individuals vs. Other Inst.	-2.09	2.96	0.15

Table 5: Relation between Informativeness and Profitability: Portfolio Sorts

This table presents portfolio analysis of trader profitability, trader profitability ratio, and trader informativeness, all estimated using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during our sample period, April to June, 2006. *Total_PL* is the average hourly profit/loss (sum of change in market value of inventory and profit/loss from round-trip transactions during the one-hour interval) estimated for each trader-stock combination. *Total_PL_Ratio* is the average of the ratio of hourly profit/loss and the capital) employed (absolute value of inventory by the trader in the stock during the one-hour interval. *Total_PL_Sharpe_Ratio* is the ratio of the difference between *Total_PL_Ratio* and the risk-free rate, expressed as a ratio of the standard deviation of *Total_PL_Ratio*. *IPRO* is the copula-based measure of informativeness calculated for each trader-stock combination. Panel A presents an in-sample analysis, where trader-stock combinations are put into deciles based on IPRO calculated using all data between April and June, 2006, and Median, Mean and t-stats are calculated and presented for *IPRO*, *Total_PL_Ratio*, and *Total_PL_Sharpe_Ratio*. Panel B presents an out-of-sample analysis, where trader-stock combinations are put into deciles based on *IPRO* calculated using the first-half of the data (between April 1st and May 14th, 2006), and Median, Mean and t-stats are calculated and presented for IPRO, *Total_PL_Ratio*, and *Total_PL_Sharpe_Ratio*, calculated using the second-half of the data (between May 15th and June 30th, 2006).

Panel A: In-Sample

`

IPRO Decile	*IPRO*			*Total_PL_Sharpe_Ratio*			*Total_PL_ratio*		
	Median	*Mean*	*t-stat*	*Median*	*Mean*	*t-stat*	*Median*	*Mean*	*t-stat*
1	-1.580	-1.713	-283.08	-0.001	-0.001	-0.06	-0.291	0.000	0.03
2	-0.958	-0.969	-593.76	-0.019	-0.012	-0.98	-0.298	-0.020	-1.55
3	-0.622	-0.625	-555.18	-0.019	-0.008	-0.60	-0.298	-0.010	-0.74
4	-0.364	-0.363	-406.47	-0.019	-0.015	-1.10	-0.288	-0.012	-0.93
5	-0.140	-0.140	-167.42	-0.055	-0.014	-1.19	-0.301	-0.015	-1.16
6	0.069	0.072	86.99	-0.059	-0.036	-2.52	-0.320	-0.030	-2.27
7	0.303	0.305	319.11	-0.048	-0.012	-0.98	-0.310	-0.009	-0.71
8	0.581	0.584	481.01	-0.023	0.013	1.09	-0.279	0.003	0.27
9	0.952	0.968	518.20	0.006	0.033	2.58	-0.269	0.037	2.80
10	1.692	1.882	225.66	0.034	0.052	3.65	-0.266	0.055	3.58

Panel B: Out-of-Sample

IPRO Decile	*IPRO*			*Total_PL_Sharpe_Ratio*			*Total_PL_ratio*		
	Median	*Mean*	*t-stat*	*Median*	*Mean*	*t-stat*	*Median*	*Mean*	*t-stat*
1	-1.516	-1.719	-153.47	-0.006	-0.004	-0.59	-0.252	-0.025	-1.40
2	-0.931	-0.944	-487.42	-0.014	0.001	0.20	-0.257	-0.014	-0.73
3	-0.656	-0.661	-522.63	-0.013	-0.053	-0.97	-0.256	-0.017	-0.95
4	-0.435	-0.438	-365.74	0.003	0.002	0.32	-0.249	0.031	1.50
5	-0.226	-0.224	-190.83	-0.007	0.004	0.67	-0.253	-0.017	-0.96
6	0.088	0.096	35.53	-0.008	0.009	2.07	-0.255	-0.006	-0.30
7	0.436	0.436	334.91	0.005	0.014	2.50	-0.247	0.019	0.99
8	0.681	0.686	458.76	0.001	0.006	0.90	-0.251	-0.009	-0.46
9	1.001	1.012	462.81	0.010	0.018	2.63	-0.249	0.009	0.49
10	1.591	1.755	179.24	0.002	0.001	0.07	-0.250	0.027	1.43

Table 6: Informativeness and Order Choice

This table presents results from Logit regressions employed to analyze the relation between traders' informativeness and order choices. The analysis is conducted using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during our sample period, April to June, 2006. *Aggressiveness Ratio* (ratio of aggressive to passive limit order volume) is calculated for each trader-stock combination over thirty-minute intervals. The dependent variable is a binary variable equal to 1 when *Aggressiveness Ratio* is greater than 0.5 for a trader in a given stock during a thirty-minute interval. *IPRO* is the copula-based measure of informativeness calculated for each trader-stock combination. *Abs Return* is the absolute value of return of a stock over the thirty-minute interval. *Bid-Ask Spread* is the average value of the ratio of the difference between the difference between the best sell price and the best buy price and the mid-quote over a thirty-minute interval. *Volume* is the total trading volume during a thirty-minute interval in a stock. *Abs Rel OIB* is the ratio of the absolute value of order imbalance over a thirty-minute interval in a stock (buy order volume – sell order volume) and *Volume*. *Pre Earnings* is a binary variable equal to 1 the day prior to the earnings' announcement in a stock. *Earnings* is a binary variable equal to 1 on the day of the earnings' announcement in a stock. *Post Earnings* is a binary variable equal to 1 on the day after the earnings' announcement in a stock. *Forex* is a binary variable equal to 1 on the 18th and 19th of May, 2006. *Open* is a binary variable equal to 1 for the first hour of daily trading. *Close* is a binary variable equal to 1 for the last hour of daily trading.

Variable	*Estimate*	*P-Value*	*Estimate*	*P-Value*	*Estimate*	*P-Value*	*Estimate*	*P-Value*
Intercept	-0.63	<.001	-0.63	<.001	-0.63	<.001	-0.63	<.001
IPRO	0.81	<.001	0.63	<.001	0.78	<.001	0.81	<.001
IPRO*Open			0.82	<.001				
IPRO*Close			0.04	0.497				
IPRO*Forex					0.62	<.001		
IPRO*Pre_Earnings							0.07	0.737
IPRO*Earnings							0.01	0.937
IPRO*Post_Earnings							-0.39	0.038
Abs Return	-0.01	<.001	-0.01	<.001	-0.01	<.001	-0.01	<.001
Bid-Ask Spread	0.36	<.001	0.36	<.001	0.36	<.001	0.36	<.001
Volume	0.03	<.001	0.03	<.001	0.03	<.001	0.03	<.001
Abs Rel OIB	0.02	<.001	0.02	<.001	0.02	<.001	0.02	<.001
Pre_Earnings							0.06	<.001
Earnings							0.01	0.359
Post_Earnings							0.08	<.001
Forex	0.01	0.2664	0.01	0.2732	0.01	0.079	0.00	0.670
Open	-0.11	<.001	-0.11	<.001	-0.11	<.001	-0.11	<.001
Close	0.19	<.001	0.19	<.001	0.19	<.001	0.19	<.001
N	5850057		5850057		5850057		5850057	
Wald Test (p-Value)	<.0001		<.0001		<.0001		<.0001	

Table 7: Informativeness and Hidden Orders

This table presents results from Logit regressions employed to analyze the relation between traders' informativeness and use of hidden orders. The analysis is conducted using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during our sample period, April to June, 2006. *Hidden Ratio* (ratio of the hidden limit order volume and the trader's total limit order volume) is calculated for each trader-stock combination over thirty-minute intervals. The dependent variable is a binary variable equal to 1 when *Hidden Ratio* is greater than 0.5 for a trader in a given stock during a thirty-minute interval. *IPRO* is the copula-based measure of informativeness calculated for each trader-stock combination. *Abs Return* is the absolute value of return of a stock over the thirty-minute interval. *Bid-Ask Spread* is the average value of the ratio of the difference between the difference between the best sell price and the best buy price and the mid-quote over a thirty-minute interval. *Volume* is the total trading volume during a thirty-minute interval in a stock. *Abs Rel OIB* is the ratio of the absolute value of order imbalance over a thirty-minute interval in a stock (buy order volume – sell order volume) and *Volume*. *Aggressiveness Ratio* (ratio of the aggressive limit order volume and the trader's total limit order volume) is calculated for each trader-stock combination over thirty-minute intervals. *Prop_Agg_Dummy* is a binary variable equal to 1 when *Aggressiveness Ratio* is greater than 0.5 for a trader in a given stock during a thirty-minute interval. *Pre Earnings* is a binary variable equal to 1 the day prior to the earnings' announcement in a stock. *Earnings* is a binary variable equal to 1 on the day of the earnings' announcement in a stock. *Post Earnings* is a binary variable equal to 1 on the day after the earnings' announcement in a stock. *Forex* is a binary variable equal to 1 on the 18th and 19th of May, 2006. *Open* is a binary variable equal to 1 for the first hour of daily trading. *Close* is a binary variable equal to 1 for the last hour of daily trading.

Variable	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value	Estimate	P-Value
Intercept	-2.83	<.001	-2.83	<.001	-2.84	<.001	-2.83	<.001	-2.83	<.001
IPRO	0.03	0.615	1.01	<.001	-0.49	<.001	0.01	0.924	0.04	0.538
IPRO*Prop_Agg_Dummy			-7.61	<.001						
IPRO*Open					1.84	<.001				
IPRO*Close					0.86	<.001				
IPRO*Forex							0.86	0.0197		
IPRO*Pre_Earnings									0.35	0.507
IPRO*Earnings									0.21	0.678
IPRO*Post_Earnings									-0.99	0.045
Abs Return	-0.08	<.001	-0.08	<.001	-0.08	<.001	-0.08	<.001	-0.08	<.001
Bid-Ask Spread	0.06	<.001	0.06	<.001	0.06	<.001	0.06	<.001	0.06	<.001
Volume	-0.06	<.001	-0.06	<.001	-0.06	<.001	-0.06	<.001	-0.06	<.001
Abs Rel OIB	-0.04	<.001	-0.04	<.001	-0.04	<.001	-0.04	<.001	-0.04	<.001
Prop_Agg_Dummy	-1.41	<.001	-1.47	<.001	-1.41	<.001	-1.41	<.001	-1.41	<.001
Pre_Earnings									-0.02	0.231
Earnings									-0.14	<.001
Post_Earnings									-0.01	0.446
Forex	-0.13	<.001	-0.13	<.001	-0.13	<.001	-0.12	<.001	-0.12	<.001
Open	-0.14	<.001	-0.14	<.001	-0.14	<.001	-0.14	<.001	-0.14	<.001
Close	-0.10	<.001	-0.10	<.001	-0.10	<.001	-0.10	<.001	-0.10	<.001
N	5850057		5850057		5850057		5850057		5850057	
Wald Test (p-Value)	<.0001		<.0001		<.0001		<.0001		<.0001	

Table 8 – Trader Categories, 2012

This table presents characteristics of Algorithmic trading in the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during the sample period, May, 2012. All market variables are calculated over 5 minute intervals. Traders are classified into three client categories by NSE (National Stock Exchange). Algo is a binary variable that identifies algorithmic messages. Institutional Traders are classified as Client 1 and Individual traders are classified as Client 3 in the data. Exchange Members are traders classified as Client 2 traders in the data. Participation is the proportion of trading volume that involves an Algorithmic trader either on the buy or the sell side. Liquidity Provision is the proportion of trading volume for which algorithmic traders provided liquidity, which is calculated based on the aggressiveness of the orders involved in the trade.

Panel A: Algo v/s Manual

Algo	Participation	Liquidity Provision
1	37.77%	16.07%
0	62.23%	33.94%

Panel B: Algo v/s Manual Across Trader Categories

Clients	Algo	Participation	Liquidity Provision
Institutional Traders	1	21.14%	10.54%
	0	12.68%	7.62%
Individual Traders	1	4.11%	2.56%
	0	36.69%	18.47%
Exchange Members	1	12.52%	2.96%
	0	12.86%	7.85%

Table 9 – IPRO and Algorithmic Traders

This table presents descriptive statistics of trader informativeness estimated using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during our sample period, May, 2012. *IPRO* is the copula-based measure of informativeness calculated for each trader category-stock combination.

Trader Type	Median	Mean	t-Value
Institutional Algo	0.34	0.19	1.34
Institutional Non-Algo	-0.17	-0.03	-0.16
Exchange Members Algo	-0.10	0.01	0.13
Exchange Members Non-Algo	-0.16	-0.09	-0.85
Individual Algo	-0.03	0.08	0.78
Individual Non-Algo	-0.20	-0.16	-1.14

Table 10: Momentum and Heterogeneity – Portfolio Analysis

This table presents results from portfolio analysis employed to analyze the relation between traders' heterogeneity and intraday momentum in returns. The analysis is conducted using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during our sample period, April to June, 2006. Stocks are sorted into three portfolios based on past 30-minute returns and independently sorted into three groups based on traders' heterogeneity measured at portfolio formation. Traders' expectation of stock prices are extracted from their limit order sub missions as explained in section XXX. Traders' heterogeneity is measured as the ratio of the inter-quartile range of traders' expectation of stock prices and the average of the expectation of stock prices. Stock returns are standardized by each stock.

		Low Heterogeneity	Moderate Heterogeneity	High Heterogeneity
Low Return	Mean	0.02	0.01	-0.05
	Median	0.00	0.00	-0.01
	t-stat	1.34	0.29	-2.41
Moderate Return	Mean	-0.02	-0.03	0.01
	Median	-0.02	-0.01	-0.01
	t-stat	-1.19	-1.93	0.55
High Return	Mean	-0.07	-0.01	0.07
	Median	-0.05	-0.04	0.01
	t-stat	-4.52	-0.88	3.60

Table 11: Momentum and Heterogeneity – Regression Analysis

This table presents estimates of Fama-MacBeth (1973) predictive cross-sectional regressions for thirty-minute stock returns. The analysis is conducted using data on the 50 stocks that make up the Standard & Poor's CNS Nifty index at the National Stock Exchange (NSE), India, during our sample period, April to June, 2006. The dependent variable is *Midquote_Return* - thirty-minute return calculated using mid-quotes. Traders' expectations of stock prices are extracted from their limit order sub missions as explained in section XXX. *IQ_Range_Ratio* is measured as the ratio of the inter-quartile range of traders' expectations of stock prices and the average of the expectations of stock prices. *Abs_Error* is the absolute value of the difference between the average of the traders' expectations of stock prices and the realized price of the stock; it is expressed as a ratio of the average of the traders' expectations of stock prices. *Abs Return* is the absolute value of return of a stock over the thirty-minute interval. *Volume* is the total trading volume during a thirty-minute interval in a stock. All variables are standardized by each stock. *P-values* are adjusted for autocorrelation following Newey-West (1987).

Variable	Estimate	P-Value
Intercept	0.06	0.005
Midquote_Return_Lag1	-0.19	0.095
Midquote_Return_Lag1* Abs_Error	-0.24	0.135
Midquote_Return_Lag1*IQ_Range_Ratio	0.31	0.126
Midquote_Return_Lag1* Abs Return	-0.02	0.624
Midquote_Return_Lag1*Volume	0.09	<.001
Volume	0.04	<.001
Abs_Error	-0.13	<.001
Abs Return	0.11	0.002
IQ_Range_Ratio	0.05	<.001
Adj. R-Square	57.49%	

References

[1] Allen, F.; S. Morris; and H. S. Shin. (2006) "Beauty Contests and Iterated Expectations in Asset Markets." Review of Financial Studies, 19, 719–752.

[2] Anand, Amber, Sugato Chakravarty, and Terry Martell. 2004. Empirical evidence on the evolution of liquidity: Choice of market versus limit orders by informed and uninformed traders. Unpublished manuscript, Syracuse University.

[3] Andrews . D. (1994) "Asympototics for Semiparametric Econometric Models via Stochastic Equicontinuity", *Econometrica*, 62, 43-72

[4] Arnold, B.C., N. Balakrishnan, and H.N. Nagaraja (1992) "*A First Course in Order Statistics*" New York: John Wiley & Sons.

[5] Angel, Jim. 1994. Limit orders versus market orders. Unpublished manuscript, Georgetown University.

[6] Bae, Kee-Hong, Hasung Jang, and Kyung S. Park. 2003. Traders' choice between limit and market orders: Evidence from NYSE stocks. Journal of Financial Markets 6:517–38.

[7] Banerjee, S.; R. Kaniel; and I. Kremer. (2009) "Price Drift as an Outcome of Differences in Higher-Order Beliefs." Review of Financial Studies

[8] Beber, Alessandro, and Cecilia Caglio. 2004. Order submission strategies and information: Empirical evidence from the NYSE. Unpublished manuscript, University of Lausanne.

[9] Biais, Bruno, Pierre Hillion, and Chester Spatt. 1995. An empirical analysis of the limit order book and the order flow in the Paris Bourse. Journal of Finance 50:1655–89.

[10] Bloomfield, Robert, Maureen O'Hara, and Gideon Saar. 2003. The "make or take" decision in an electronic market: Evidence on the evolution of liquidity. Unpublished manuscript, New York University, Stern School of Business.

[11] Chakravarty, Sugato, and Craig Holden. 1995. An integrated model of market and limit orders. Journal of Financial Intermediation 4:213–41.

[12] Chan, Louis K. C., and Josef Lakonishok. 1993. Institutional trades and stock price behavior. Journal of Financial Economics 33:173–99.

[13] Chen, X and Y.Fan (2002) "Estimation of Copula Based Time Series Modesl" Working Paper. NYU

[14] Diether, K. B.; C. J. Malloy; and A. Scherbina. (2002) "Differences of Opinion and the Cross Section of Stock Returns." Journal of Finance, 57, 2113–2141.

[15] Easley, David, N. Kiefer, and Maureen O'Hara. 1996. Cream-skimming or profit-sharing? The curious role of purchased order flow. Journal of Finance 51:811–33.

[16] Easley, David, N. Kiefer, Maureen O'Hara, and J. Paperman. 1996. Liquidity, information, and less-frequently traded stocks. Journal of Finance 51:1405–36.

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

[18] ———. 1991. Order form and information in security markets. Journal of Finance 47:577–606.

[19] Ellul, Andrew, Craig W. Holden, Pankaj Jain, and Robert H. Jennings. 2003. Determinants of order choice on the New York Stock Exchange. Unpublished manuscript, Indiana University.

[20] Embrechts, P. A. Mcneil and D. Straumann (2001) "Correlation and Dependence in Risk Management: Properties and Pitfalls" Working Paper.

[21] Foucault, Thierry. 1999. Order flow composition and trading costs in a dynamic limit ordermarket. Journal of Financial Markets 2:99–134.

[22] Foucault, Thierry, Ohad Kadan, and Eugene Kandel. 2002. Limit order book as a market for liquidity. Unpublished manuscript, Washington University.

[23] Genset, C.(1987) "Frank's Family of Bivariate Distributions," *Biometrika*, 74 (3), 549-555

[24] Genset, C. (1987) "A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions," *Biometrika*, 82(3), 543-552.

[25] Genset, C.and L.P. Rivest (1993) “Statistical Inference Procedure for Bivariate Archimedean Copulas," *Journal of the American Statistical Association*, 88 (423), 1034-1043.

[26] Genset, C. K. Ghoudi, and L.P. Rivest (1995) “A Semiparametric Estimation Procedure of Dependence Parameters in Multivariate Families of Distributions," *Biometrika*, 82(3), 543-552.

[27] Glosten, Lawrence R. 1994. Is the electronic open limit order book inevitable? Journal of Finance 49:1127–61.

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

[29] Hamao, Yasushi, and Joel Hasbrouck. 1995. Securities trading in the absence of dealers: Trades and quotes on the Tokyo Stock Exchange. Review of Financial Studies 8:849–78.

[30] Handa, Puneet, and Robert A. Schwartz. 1996. Limit order trading. Journal of Finance 51:1835–61.

[31] Harris, Lawrence. 1998. Optimal dynamic order submission strategies in some stylized trading problems. Financial Markets, Institutions and Instruments 7:1–76.

[32] Harris, Lawrence, and Joel Hasbrouck. 1996. Market vs. limit orders: The SuperDOT evidence of order submission strategy. Journal of Financial and Quantitative Analysis 31:213–33.

[33] Hasbrouck, Joel. 1996. Order characteristics and stock price evolution: An application to program trading. Journal of Financial Economics 46:179–208.

[34] Hasbrouck, Joel, and George Sofianos. 1993. The trades of market makers: An empirical analysis of NYSE specialists. Journal of Finance 48:1565–93.

[35] Hasbrouck, Joel, and Deborah Sosebee. 1992. Orders, trades, reports and quotes at the NYSE. Unpublished manuscript, New York Stock Exchange.

[36] Hollifield, Burton, Robert Miller, Patrick Sandas, and J. Slive. 2002. Liquidity supply and demand in limit order markets. Unpublished manuscript, University of Pennsylvania, Wharton School.

[37] Hong, Harrison, and Jiang Wang. 2000. Trading and returns under periodic market closures. Journal of Finance 55:297–354.

[38] Huang, Roger D., and Hans R. Stoll. 1996. Dealer vs. auction markets: A paired comparison of execution costs on NASDAQ and the NYSE. Journal of Financial Economics 43:313–57.

[39] Kavajecz, Kenneth, and Elizabeth R. Odders-White. 2001. An examination of changes in specialists' posted price schedules. Review of Financial Studies 14:681–704.

[40] Joe (1997) " Multivariate Models and Dependence Concepts"

[41] Kaniel, Ron and Hong Liu (2006) " So What Orders Do Informed Traders Use", Journal of Business, vol. 79, no. 4, pp 1867-1913.

[42] Nelson R. (1999) "*Introduction to Copulas*" Springer, NY.

[43] Rao, B.L.S.P. (1992) *"Identifiability in Stochastic Models: Characterization of Probability Distributions"* Academic Press, New York.

[44] Scheinkman, Jose and Wei Xiong (2002) *"Heterogeneous Beliefs, Speculation and Trading in Financial Markets"*

[45] Sklar, A. (1959) Fonctions De repartition a n dimensions et lers marges. *Publ. Inst. Statist. Univ.* Paris., 8. 229-31.