Do Investors Benefit from More Transparency?
An Asset Allocation Perspective*

Ingmar Nolte† Richard Payne‡ Michalis Vasios§

Abstract

A current debate in finance concerns the transparency in financial markets and the disclosure of counterparty identity information. We use a simple mean-variance framework and data from Helsinki Stock Exchange to explore the asset allocation implications of market transparency. We find that broker identity conveys information that is economically significant. A mean-variance investor can benefit remarkably, up to 36% (annualized) percentage points for the most parsimonious forecasting model, from observing the order flow of a post-trade non-anonymous market. This result suggests that market transparency yields positive economic value. A second result is the substantial variation in the information content of order flow at the broker level. We show that the predictive power of broker customer order flow can be attributed to observable broker-specific characteristics: market share, daily volume, investment style and degree of sophistication.

Keywords: Market Transparency, Asset Allocation, Broker Heterogeneity, Customer Order Flow.

This Version: December 11, 2011

*We would like to thank Nicholas Barberis, Pasquale Della Corte, Charles Jones, Heracles Polemarchakis, and participants at Warwick Business School, the 2011 Trans-Atlantic Doctoral Conference (London Business School), the 2011 Frontiers of Finance Conference (Warwick Business School), the 2011 High Frequency Research Training Workshop (The Marie Curie Risk Network), and the 12th Symposium on Finance, Banking, and Insurance (Karlsruhe Institute of Technology) for helpful comments.

†Warwick Business School, Finance Group, Financial Econometrics Research Centre (FERC), University of Warwick, CV4 7AL, Coventry, UK. Phone +44 2476 572 838, Fax -23779, email: Ingmar.Nolte@wbs.ac.uk.

‡Cass Business School, 106 Bunhill Row, London EC1Y 8TZ, UK. Phone +44 207 040 4114, email: Richard.Payne.1@city.ac.uk.

§Corresponding author: Warwick Business School, Finance Group, Financial Econometrics Research Centre (FERC), University of Warwick, CV4 7AL, Coventry, UK. Phone +44 7942 468 063, email: michalis.vasios.09@mail.wbs.ac.uk.

An ongoing debate in finance among researchers, practitioners, and policy makers is about transparency in financial markets and the disclosure of counterparty identity information. Over the last decade there has been a global trend towards anonymity, which turned most major markets opaque (e.g., Euronext Paris in 2001, London Stock Exchange's SETS in 2001, Italian Stock Exchange (Borsa Italiana) in 2004). In recent years, markets have moved to even less-transparent structures with the emergence of the so-called dark pools[1], which are alternative trading systems (ATSs) that do not publicly display quotes.

While the tendency to less transparent and more anonymous markets can be justified to some extent by the evidence that anonymity supports the provision of liquidity[2], it is unclear whether the lack of transparency has any side-effects. Early theoretical evidence (e.g., Forster & George (1992), Benveniste, Marcus & Wilhelm (1992) and Rindi (2008)) suggests that the degree of transparency affects the information content of order flow, the distribution of trading profits, and inevitably market efficiency. The 2008 financial crisis has turned the interest of policy makers on these issues and triggered a new trend towards transparency. For instance, the SEC proposed[3] in 2009 the enhancement of transparency for dark pools both pre-trade, with the public display of best quotes; and post-trade, with the disclosure of the identity[4] of the ATS or other broker-dealer that reports the trades. Their motivations was

> " (...) to ensure that publicly available prices fully reflect overall supply and demand, equip the investing public with tools to make better investment decisions, increase the perception of fairness that is necessary for the healthy functioning of the national market system, and, as a result, enhance public confidence in the securities markets."

In the proposed rule, the SEC acknowledges the risk of the disclosure of identities to cause inappropriate information leakage on dark pools' customer flow recognizing that the broker-dealer identity does convey information. The SEC also acknowledges that

[1] According to the Securities and Exchange Commission (SEC), as for September 2009, 32 dark pools in the United States accounted for 7.9% of total equity trading volume; Securities and Exchange Commission; Concept Regulation of Non-Public Trading Interest; Proposed Rule, November 23, 2009, Federal Register 74(224), 61207-61238.

[2] For instance, Comerton-Forde, Frino & Mollica (2005), Foucault, Moinas & Theissen (2007), and Comerton-Forde & Tang (2009) find that market liquidity increases significantly in anonymous markets. Similarly, Friederich & Payne (2011) show that post-trade anonymity greatly improved liquidity on the London Stock Exchange.

[3] See *supra* note 1.

[4] ATSs can register either as national securities exchanges or as broker-dealers that are members of FINRA. While for registered exchanges their identity is attached on trade reports, for broker-dealers their identity is not revealed to the public.

investors have the right to have access to this information in order, among others, to improve their decision making. Nonetheless, the information content of order flow in a non-anonymous market and the asset allocation implications of the degree of transparency are not fully understood to date.

In this paper we explore the information content of order flow in a post-trade non-anonymous market, the Helsinki Stock Exchange (HEX), and in particular we test the hypothesis that investors benefit from more transparency through the observation of trading activity at the broker firm level. Does broker identity convey information? Do investors make better decisions in a non-anonymous market?

We test this hypothesis by building dynamic mean-variance portfolios, in the fashion of Fleming, Kirby & Ostdiek (2001) and Marquering & Verbeek (2004), that use the broker order flow information to predict next period's returns and by comparing their performance against a benchmark portfolio that disregards the broker identity. In our model we capture trading activity and information by the *order flow* measure[5] of Chordia, Roll & Subrahmanyam (2002) and Evans & Lyons (2002), which reflects the buy minus the sell market pressure in a pre-specified interval, which in our setting is a day. Based on the performance difference against the benchmark portfolio we measure the predictive power of broker identity on the one hand, and provide an indication on the value of the missing information when markets switch to less transparent structures - the economic value of market transparency - on the other hand. In addition, our approach allows us to compare the information content of order flow of different brokers and examine whether the dynamics of information generation and aggregation at the broker firm level are due to broker-specific characteristics: size, investment style and degree of sophistication.

Our empirical analysis suggests that the broker identity conveys information that is economically significant. More specifically, we show that an investor with mean-variance preferences can greatly improve (up to 36% (annualized) percentage points for the most parsimonious forecasting model) his portfolio performance[6] by observing brokers' customer order flow activity and using this information to update his beliefs and rebalance his portfolio at a daily frequency. We also find that the informativeness of order flow is merely driven by aggressive trading, while the observation of liquidity supplying either does not improve or improves only marginally portfolio performance.

[5]Kyle (1985) and Glosten & Milgrom (1985) provide the theoretical foundations for the information content of order flow. Empirically, evidence from Hasbrouck (1991), Evans & Lyons (2002), Payne (2003), Love & Payne (2008), and Rime, Sarno & Sojli (2010) among others, demonstrates the important role of order flow on the transmission of private and public information into prices.

[6]We measure performance by the manipulation-proof performance measure (MPPM) of Goetzmann, Ingersoll, Spiegel & Welch (2007).

Since the performance of dynamic investment strategies is sensitive to transaction costs, we control for that and find their role to be minor or in some cases even supportive to our results.

A second result is the variation of the information content of order flow at the broker firm level. While we find evidence that the broker identity reveals valuable information, this does not hold for all brokers. Given this finding, the improvement on investment decision making is conditional on investors ability to identify the better informed brokers. If this identification is impossible, then being in a transparent or an opaque market would not make any difference from an asset allocation perspective. Our analysis, however, shows that the predictive power of order flow can be attributed[7] to observable broker-specific characteristics.

We start by exploring the relation between broker identity predictive power and market share. Previous literature[8] suggests that market leaders have better information. To test if this assertion holds in HEX, we compute the performance difference of a portfolio that tracks large brokers and a portfolio that tracks small brokers. A striking result is that the portfolio of large brokers significantly underperforms the portfolio of small brokers. Moreover, the order flow information of large brokers has on average no predictive power at the daily horizon. Sapp (2002) finds a similar result in a foreign exchange market context. We argue that market leaders attract very heterogeneous investors, whose aggregation at the broker level is uninformative for other investors. We test this hypothesis by calculating daily correlated trading statistics. Our analysis, then, shows that the customer order flow of large brokers is characterized by a high dispersion of beliefs (low correlated trading) that hides any meaningful signal and might explain the low informativeness of their order flow.

Although the identity of brokers that trade regularly high volumes has on average no predictive power, heavy volume on a particular day does convey information. We show that when investors observe the order flow of brokers that traded more actively the previous day, they make better investment decisions and improve their portfolio performance. This finding suggests that active trading is the transmission mechanism of new information into prices. As we show, the transmission of new information lasts for one day only.

Next we investigate whether there is a linkage between the information content of order flow and the investment style of the average customer of brokers. Our motivation

[7] In a different setting, Linnainmaa & Saar (2010) show that broker identities can be used to extract meaningful signals about the types of investors who initiate trades.

[8] See e.g., Goodhart (1988), Lyons (1997), and Peiers (1997).

is the extensive literature[9] that suggests that sophisticated market participants exhibit a momentum investment style; buy past winners and sell past losers. We use the framework of Grinblatt & Keloharju (2000) to characterize brokers as momentum or contrarian. We find that the identity of momentum brokers has strong predictive power for future returns. In contrast, the customer order flow of momentum brokers has no predictive power. The performance difference between the momentum and the contrarian portfolio is larger when investors observe both aggressive and passive trading activity. We explore whether this finding is driven by the degree of sophistication, as the previous literature has suggested. We measure broker stock picking ability, which we regard as a proxy for sophistication and find that the more sophisticated brokers exhibit a momentum trading behavior, while the less sophisticated brokers are contrarian.

Overall, our results indicate that broker identity conveys information that is not only economically significant, but also exploitable due to the link between the broker identity predictive power and broker-specific characteristics.

Our results contribute to the literature on market transparency by providing new evidence on the asset allocation implications of the degree of market transparency. The switching to a non-anonymous market structure equips investors with an additional tool, the broker identity, in their effort to optimize their investment decisions; market transparency yields positive economic value to mean-variance investors. Although this tool is available to every investor, it is reasonable to assume that it will be particularly beneficial to less sophisticated investors, who at almost no cost can have a more complete and accurate picture of trading activity at the broker level. From a distribution of trading profits point of view, it is also reasonable to argue that in transparent markets sophisticated investors will have to share their trading profits with the more "naive" investors. Forster & George (1992) and Rindi (2008) provide supportive theoretical evidence to these arguments. In contrast, switching to a less transparent market will allow sophisticated investors to trade, more likely, at the expense of less sophisticated traders. To the extent that the perception of fairness and the public confidence in securities markets affects policy makers decisions, our results justify the recent trend[10] in US and Europe towards the enhancement of transparency in securities markets.

We also shed light on the determinants of the information content of broker cus-

[9]See e.g., Grinblatt, Titman & Wermers (1995a), Grinblatt & Keloharju (2000), Goetzmann & Massa (2002). Bloomfield, Tayler & Zhou (2009) show that short-term momentum is mainly caused by informed traders.

[10]See supra note 1 for the proposed policy in US. European Union is revising the Markets in Financial Instruments Directive (Directive 2004/39/EC) aiming "at establishing a safer, sounder, more transparent, and more responsible financial system that works for the economy and society as a whole."

tomer order flow. Our results suggest that brokers play an important role in acquiring trading activity and information. Their distinct characteristics affects this “aggregation process” in a different way. For example, market leaders aggregate very uncorrelated trading activity that makes their order flow less informative. In contrast, small brokers’ persistent trading is a better predictor of future market developments. Similarly, we find a connection between the investment style of the average customer of brokers and the predictive power of their order flow. The link between broker heterogeneity and information aggregation is new in the literature, since previous papers[11] treat brokers as a special case of informed investors that trade on their own interest. Our results imply a more active role of brokers in the aggregation of information and its transmission into prices.

The rest of the paper is organized as follows. The next section describes the data set we use. Section 2 presents the empirical framework. In Section 3, we discuss the empirical results of the mean-variance analysis. In Section 4, we investigate the determinants of the information content of order flow and broker heterogeneity. Section 5 concludes.

1 Data and Summary Statistics

We use equity intraday data from Helsinki Stock Exchange (HEX), which are provided by Bloomberg. HEX is a part of NASDAQ OMX Group since 2007, when NASDAQ acquired OMX. As advertised on their site, the NASDAQ OMX Group is the world’s largest exchange company. NASDAQ OMX Nordic describes the common offering from NASDAQ OMX exchanges in Helsinki, Copenhagen, Stockholm, Iceland, Tallinn, Riga, and Vilnius. These exchanges use a common trading platform, which allows for cross-border trading and settlement, and cross membership. According to NASDAQ OMX Nordic site, its trading platform offers access to more than 80 percent of the exchange trading in the Nordic countries.

Our data set begins at 8am (GMT) on Monday 29^{th} March 2010 and ends at 4:30pm on Monday 28^{th} February 2011; this amounts to 210 trading days. In our analysis, we consider the 15 most liquid (in terms of turnover) stocks of HEX25 index, in order to circumvent problems arising from the low number of transactions of some brokers. Table 2 reports the summary statistics of the 15 stocks in the total sample period, the initial estimation window, and the out-of-sample period. Every day there

[11] For instance, Roell (1990), Fishman & Longstaff (1992) and Pagano & Roell (1993) examine the effects of dual trading and front running. Forster & George (1992) and Anand & Subrahmanyam (2008) focus on broker trading on their own account.

are 4 regular trading sessions: opening (7am-8am), continuous trading (8am-4:25pm), closing (4:25pm-4:30pm) and after market (4:30pm-7am). We restrict the empirical analysis to the continuous trading session. In NASDAQ OMX Nordic markets, the broker code is public information during the continuous trading session.

The raw data contains information on 7 items. Table 3 represents 5 seconds of the transaction data of Nokia (NOK1V). The first two columns are the date and time expressed as month/day/year and hour:minute:second, respectively. The third column is the type of the transaction, which can be "Best Bid", "Best Ask" or "Trade". The next two columns are the price (in Euros) and the size of the transaction. The last two columns are the Broker Buy Code and the Broker Sell Code. In Table 1 there is the list of brokers[12] in HEX. The counterparty identity is available only for transactions of type "Trade".

The data record does not give information on the direction of the trade. However, the availability of best bid and best ask quotes, as well as their time stamp enables us to identify which broker initiated the trade and, thus, the direction of the trade. This identification is an important element in our empirical analysis, since it allows us to disaggregate the data and construct distinct order flow measures for every broker.

2 Empirical Framework

2.1 The Formation of the Dynamic Mean-Variance Portfolios

Our empirical analysis relies on the formation of dynamic mean-variance portfolios. Our investment scenario considers an investor with mean-variance preferences, who allocates his wealth across the 15 most liquid stocks of HEX25 index and the risk-free asset. Rebalancing is daily and conditional on the observation of the previous day's trading activity, which is captured by order flow measures. HEX being a non-anonymous market allows investors to observe not only the aggregate market order flow, but also the customer order flow of brokers. Either order flow is an input to

[12] We drop brokers that do not trade in all stocks, are acquired by other brokers, are not members in all the sample period, and trade very infrequently (initiate (on average, across stocks) less that 10 trades every day and are active (on average, across stocks) in less than half days of our sample period) from the list of brokers. We do that to deal with the singular matrices when we estimate the order flow models (Section 2), and to increase the power of the LSV statistic and the buy ratios (Section 4). Results remain qualitatively the same if instead we use all brokers.

investor's optimization problem, which is given by:

$$\begin{aligned} \max_{w_t^j} \ \mu_{p,t+1|t} &= (w_t^j)'\mu^j_{N,t+1|t} + (\iota - (w_t^j)'\iota)R_t^f \\ s.t. \ \sigma_p^2 &= (w_t^j)'\Sigma_{t+1|t}w_t^j, \qquad t = 1, \ldots, T, \end{aligned} \tag{1}$$

where $j = 1, \ldots, J$ identifies broker^j, w_t^j is the $N \times 1$ vector of portfolio weights; $\mu_{p,t+1|t}$ is the conditional expected portfolio return; σ_p is the target volatility; $\Sigma_{t+1|t}$ is the $N \times N$ variance-covariance matrix of the risky assets; R_t^f is the risk free rate; and $\mu^j_{N,t+1|t}$ is the $N \times 1$ vector of expected returns of the risky assets conditional on the order flow information of broker^j, $\mu^j_{N,t+1|t} = E[R_{t+1}|\Im_t^j]$. The solution to this constrained maximization problem yields,

$$w_t^j = \frac{\sigma_p \ (\Sigma_{t+1|t})^{-1} \ (\mu^j_{N,t+1|t} - \iota R_t^f)}{\sqrt{(\mu^j_{N,t+1|t} - \iota R_t^f)' \ (\Sigma_{t+1|t})^{-1} \ (\mu^j_{N,t+1|t} - \iota R_t^f)}}. \tag{2}$$

These are the weights for the risky assets at each rebalancing time interval. The investment in the risk free asset is equal to $1 - (w_t^j)' \ \iota$. Then, the period $t+1$ gross return on the investor's portfolio is given by $1 + (w_t^j)'R_{t+1} + (1 - (w_t^j)'\iota)R_f$.

A key element in Equation 2 is the vector of conditional expected returns of risky assets. We presume that the information set of the aggregate market and that of brokers differ. We approximate these information sets by using transaction data to compute order flow measures. There is an extensive literature[13] on order flow and how it can impact returns not only through short-term liquidity and inventory effects, but also because it conveys information. Our methodological contribution to this literature is the disaggregation of the order flow measure at the aggregate market and the broker firm level. That means that for every stock, we have as many conditional expected return estimates as the number of brokers plus the aggregate market estimate.

In our analysis, we use 2 order flow specifications. The first one is the standard *order flow* (OF_t^j) measure of Chordia et al. (2002) and Evans & Lyons (2002), defined as the daily buyer-initiated volume minus the seller-initiated volume. This measure captures aggressive trading which is considered to transmit new information into prices. To see whether liquidity-supplying (passive trading) conveys information too, we use a second specification[14] ($VolOF_t^j$), defined as the total buy volume minus the total sell

[13] See e.g., Hasbrouck (1991), Chordia et al. (2002), Evans & Lyons (2002), Easley, Hvidkjaer & O'Hara (2002), Payne (2003), Pasquariello & Vega (2007), Evans & Lyons (2008), Berger, Chaboud, Chernenko, Howorka & Wright (2008), Love & Payne (2008), Nolte & Nolte (2010), Rime et al. (2010).

[14] To clarify things, the difference between OF_t^j and $VolOF_t^j$ is that the former uses volume initiated, while the latter uses volume executed (passive plus aggressive volume). We are able to calculate

volume executed by broker^j.

Building on these order flow measures, we use 4 parsimonious models to compute one-day-ahead estimates of stock returns. The first model (M1) is a pure order flow model:

$$R_{t+1}^i = \alpha + \beta OF_t^{ij} + \epsilon_{t+1}, \qquad i = 1, \ldots, 15, \tag{3}$$

where j identifies broker^j, R_{t+1}^i is the return of stock i, OF_t^j is the order flow measure of broker^j, β is a coefficient, α is a constant, and ϵ_t the error term. To capture passive trading activity, Model 2 (M2) uses the second order flow measure, $VolOF_t^j$, as an additional variable:

$$R_{t+1}^i = \alpha + \beta OF_t^{ij} + \gamma VolOF_t^{ij} + \epsilon_{t+1}, \tag{4}$$

The last two models, model 3 (M3) and model 4 (M4), extent models 1 and 2, respectively, with the introduction of the market return (HEX25 index), MKT_t, that captures the market-wise activity at time t:

$$R_{t+1}^i = \alpha + \beta OF_t^{ij} + \gamma MKT_t + \epsilon_{t+1}, \tag{5}$$

and

$$R_{t+1}^i = \alpha + \beta OF_t^{ij} + \gamma VolOF_t^{ij} + \delta MKT_t + \epsilon_{t+1}. \tag{6}$$

2.2 Performance Measures

The next step in our analysis is to measure the performance of the mean-variance portfolios. We use an economic evaluation approach and two criteria; a traditional performance measure, the Sharpe ratio, and a utility-based measure, the manipulation-proof performance measure (MPPM) of Goetzmann et al. (2007). The first economic criterion, the ex-post Sharpe ratio (SR), is defined as:

$$SR^j = \frac{\overline{r_p^j - r_f}}{\sigma_p^j}, \tag{7}$$

where the nominator is the average (annualized) excess portfolio return and the denominator is the portfolio's (annualized) standard deviation. Intuitively, the Sharpe ratio measures the risk-adjusted annualized portfolio's returns.

$VolOF_t^j$, because our dataset contains the identity of the broker that bought and sold in every transaction. By construction, this order flow definition is zero for the aggregate market; the daily buy volume always equals the daily sell volume.

The second economic criterion, MPPM, is defined as:

$$MPPM^j = \frac{1}{(1-\gamma)\Delta t} \ln \left[\frac{1}{(T-1)} \sum_{t=1}^{T-1} \left(\frac{R^j_{p,t+1}}{R^f_{t+1}} \right)^{1-\gamma} \right], \tag{8}$$

where $R^j_{p,t+1}$ is the gross portfolio return obtained when using broker's j order flow to forecast expected returns, R^f_{t+1} is the gross risk free return, Δt is the one day interval, and γ can be seen as the investor's relative risk aversion coefficient. $MPPM^j$ can be interpreted as the annualized continuously compounded excess return certainty equivalent of the broker^j portfolio. The advantage of this economic measure is that it does not require an assumption of the investor's utility function, and it is robust to the distribution of the portfolio returns.

Our interest lies on the performance differences rather than on the performance of the mean-variance portfolios per se. We use the portfolio that uses the aggregate market order flow (ANON), which is the one that disregards the broker identity, as benchmark. The performance difference against the ANON portfolio allows us to measure the predictive power of the customer order flow of brokers. If broker identity contains no information, this difference should be zero. In contrast, a positive performance difference[15] will show the predictive power of broker identity and will unveil the positive economic value of market transparency.

We test this hypothesis by calculating the following performance difference measure:

$$\Theta^j = MPPM^j - MPPM^{ANON}. \tag{9}$$

Θ^j enables us to compare competitive dynamic investment strategies. Intuitively, it is the fee that a mean-variance investor is willing to pay to switch from the benchmark asset allocation strategy to the strategy under investigation. A positive Θ^j will mean that the investor will be better-off using broker^j order flow information than using the aggregate market order flow information.

There is a number of papers that use utility-based measures to determine the economic value of a dynamic strategy versus a passive strategy. For instance, Fleming et al. (2001) investigate the economic value of volatility timing or Marquering &

[15] Under rational expectations and efficient markets, the performance difference should be close to zero even if the broker identity conveys information: prices incorporate all available public information. Nonetheless, we do document positive performance differences and under the efficient markets viewpoint our results are conservative estimates of the predictive power of brokers' customers order flow. Intuitively, in a market where some investors have privileged access to customer order flow (e.g. in a anonymous market in which the broker order flow information is private) the value of market transparency will be even larger.

Verbeek (2004) analyze the economic value of predicting both stock index returns and volatility. While we follow a similar approach, a critical difference is the fact that our benchmark strategy is not passive, but dynamic. It is the information sets captured by order flow models that differ not the style of investment strategies.

3 Empirical Results

In this section, we measure the predictive power of the broker identity. The investment scenario is based on an investor with γ=6 coefficient of relative risk aversion, who maximizes his expected portfolio return subject to an annual target volatility of σ_p=10%. Our choice of σ_p and γ is consistent with the previous relevant literature (see e.g., Rime et al. (2010) and Della Corte, Sarno & Sestieri (2010)), while for different values of σ_p and γ results remain qualitatively unchanged. Our choice of risk free rate is the ECEUR1M rate, which is available at a daily basis. We use the order flow models M1-M4, described in Section 2 to predict returns and rebalance portfolio weights. This recursive out-of-sample regression estimation is based on a window of expanding size that means that the investor uses all available historical information at time t to update his beliefs and optimize his asset allocation. The initial estimation window is 03/30/2010–08/09/2010 (86 days or 40% of the sample period) and the portfolio formation and rebalancing spans from 08/10/2010 to 02/28/2011 (124 days or 60% of our sample period). Results remain qualitatively the same if instead we split the sample period in two equal windows. We compute the variance-covariance matrix of the risky assets recursively using at each point t daily data of the previous one year to forecast volatility at t+1.

3.1 Preliminary Analysis

Before we proceed to the recursive out-of-sample estimation we present some preliminary results to obtain an indication of the statistical performance of the order flow models in the initial estimation window. To save space we present results only for model M1, however, results for the other models are available upon request.

Table 4 reports the heteroskedasticity and autocorrelation corrected (Newey-West) t-statistics of the lagged order flow variable of model M1. In bold are the statistical significant t-statistics (p-value<10%). The significance of the coefficient of the order flow variable varies across brokers and stocks. On the one hand, there are brokers like SHB or SWB with many strongly significant coefficients, while on the other hand there are brokers like NIP and NON with barely any significant coefficient. The sign of the order flow coefficients is positive for half of the estimates, while for almost one third of the brokers, the majority of their coefficients are positive. The positive coefficients indicate that the order flow and next day's return are positively related; a buy (sell) pressure on a particular trading day predicts a price increase (decrease) the next trading day, but it remains to see whether this is reflected in the portfolios' performances.

In Table 5, we present the economic evaluation of the mean-variance portfolios' performances in the initial estimation window. We want to stress that this is an in-sample estimation, since the order flow models' coefficients are estimated only once using the first 86 days of our dataset. In short, the results show that investors can improve their portfolio performance when they observe brokers' customer order flow compared to the benchmark case, which is the portfolio that disregards the broker identity. The number of positive Θ ranges from 18 for Model 1 to 28 for Model 4. However, not all of these Θ are statistical significant, with the best model being Model 4 with 9 statistical significant Θ (p-value$<$ 10%). Among the brokers that perform well across all 4 models are DBL, DDB, JPM, NRD, and SWB. As for the Sharpe ratios, they are high across all brokers and models, which is expected as these are in-sample calculations with daily rebalancing. Their magnitude is consistent with other papers that use order flow models with daily rebalancing. For instance, Rime et al. (2010) find in-sample Sharpe ratios that range from 5.79 to 7.05.

The results in this section support our hypothesis about the positive economic value of market transparency. However, the real test lies on the out-of-sample evaluation of the recursive forecasts and the performance of the mean-variance portfolios that follow in the next section.

3.2 Does Broker Identity Convey Information?

In this section, we test the hypothesis that the broker identity conveys information and that investors can benefit from transparency. Our analysis is based on an out-of-sample recursive regression estimation.

Table 6 presents the economic evaluation of the mean-variance portfolios. For the majority of brokers, Sharpe ratios are large, positive and greater than the Sharpe ratio of the ANON portfolio, which is negative. Same is true for Θs, which are significant for a large fraction of brokers. The reported p-values are below at least 10% (5%) for 11 (7), 7 (2), 11 (7), and 7 (2) brokers for Models 1 to 4, respectively. Among the brokers that perform well across all 4 models are CAR, NRD, RBN, SHB and SWB. One interpretation of a positive Θ^j is that it measures the maximum performance fee the mean-variance investor is willing to pay to switch from the ANON portfolio, which is the one that disregards the broker identity, to the portfolio that tracks broker^j customer order flow.

Our results, also, show that most of the predictive power comes from the aggressive trading activity. The passive trading activity, $VolOF_t^j$, and the market index, MKT_t, either enhance marginally or decrease the magnitude and significance of Θs. This result is not surprising, since a stylized fact in market microstructure literature is the connection of liquidity supplying with noninformative trading, which adds noise in financial markets (see e.g. Kyle (1985)). Under this perspective, the observation of brokers' customers passive trading activity should have a marginal effect on the informativeness of brokers' order flow.

The evidence to this point suggests that the broker identity conveys information. This result is in line with the paper of Linnainmaa & Saar (2010), who find that the broker identity can be used in HEX to extract a meaningful signal about the quality of the investors who initiate the trades. Forster & George (1992) provide theoretical support to this argument. We further find that this signal is economically significant from an asset allocation perspective: market transparency yields positive economic value. Investors can erase the noise in the aggregate market and greatly improve their investment decision making, up to 36% (annualized) percentage points for Model 1, by observing brokers' customer (aggressive) order flow. Our focus is on the dynamic allocation across the risky assets and the risk-free security.

Another finding from Table 6 is the strong heterogeneity of brokers. Sharpe ratios range from -1.61 to 2.93, -2.04 to 2.90, -2.03 to 2.28, and -2.61 to 2.68 for Models 1 to 4, respectively. Θ varies too: from -10 to 36, -15 to 38, -7 to 38, and -13 to 43 for Models 1 to 4, respectively. It is not clear why the predictive power of identity is strong for some brokers and zero for some others. From a practical point of view, it is important to know what drives this heterogeneity in order to understand the dynamics of information generation and aggregation at the broker level. Before we elaborate upon this issue, we discuss the role of transaction costs on the performance of the mean-variance portfolios.

3.3 Is Predictive Power Robust to Transaction Costs?

In the literature evaluating the performance of dynamic investment strategies, transaction costs play a key role. Moreover, the more frequent the portfolio rebalancing is the more significant is the impact of transaction costs on determining returns and evaluating the overall performance of portfolios. However, in our analysis there are reasons to believe that the impact of transaction costs will be limited. First, the benchmark portfolio in our framework is another dynamic investment strategy, thus, in the presence of transaction costs its performance will be affected too. Second, as can be seen in Table 7, the order flow measures of brokers are correlated to some extent with the aggregate market order flow measure, which suggests that the daily turnover of the competitive portfolios should be of the same magnitude.

To examine this issue more carefully, we repeat the formation, rebalancing and evaluation of the mean-variance portfolios in the presence of transaction costs (TC). Tables 8 and Table 9 present the Θ performance measures for three levels of transaction costs ($TC = 0bps$, $TC = 10bps$, and $TC = 30bps$) in the initial and out-of-sample period, respectively. For instance, for Model 1 in Table 9 as we move from the case of $TC = 0bps$ to the case of $TC = 30bps$, brokers' Θ become larger and more significant. A similar pattern emerges for the other 3 models. To sum up, our results are robust to the presence of transaction costs; transaction costs either play a minor or even a supportive role.

4 The Determinants of the Information Content of Broker Customer Order Flow

Our analysis suggests that the decomposition of information and trading activity in a non-anonymous market helps investors to make better investment decisions. However, this result depends on the ability of investors to select the brokers with the most informative customer order flow. Here, we explore the determinants of the information content of brokers' order flow. We test several hypotheses and provide evidence that the predictive power of order flow at the broker level can be attributed to observable broker-specific characteristics.

4.1 Does Market Leaders' Identity Signal Better Information?

We start by exploring the role of brokers' market share. The simple intuition underlying the market share hypothesis is as follows. Investors pay close attention to the trading activity of market leaders. Who wouldn't take into account the trades of Goldman Sachs or other big players? Several papers, provide the reason why; banks with large market share have the best information (see e.g., Goodhart (1988), Lyons (1997), and Peiers (1997)). This must be true especially in markets, in which only a few brokers control most of the trading activity.

As shown in Table 10, HEX belong to this category. We calculate market share with respect to the average daily volume initiated or executed by each broker across the 15 most liquid stocks of HEX25 index. In addition, we consider the fraction of average daily aggressive trading to average daily volume, which should be seen as a measure of the "trading aggressiveness" of brokers. Table 10 shows that the brokerage industry in HEX is highly concentrated; the top 5 brokers initiate almost 40% of the trading and execute 35% of the volume. The aggressiveness of brokers varies; for some brokers like CDG, NIP, and SGP more than 68% of their trading is aggressive, while for some other brokers like SHB, MSI, and UN this fraction falls to one-third.

To test the market share hypothesis, we split brokers into quartiles with respect to the three criteria of Table 10, construct daily average order flow measure series of the top and bottom quartile, and repeat the formation, rebalancing, and evaluation of the two mean-variance portfolios following the steps described in Section 2. We report the performance difference $\Delta\Theta$ that is defined as:

$$\Delta\Theta = \Theta_{Q4} - \Theta_{Q1}, \tag{10}$$

where Θ_{Q4} is the MPPM of top quartile and Θ_{Q1} is the MPPM of the bottom quartile. A positive and significant $\Delta\Theta$ will show that large brokers have more informative order flow than small brokers.

Table 11 presents the performance differences between the large and small brokers, and the associated p-values. Surprisingly, we find evidence that rejects the market share hypothesis. Clearly, the results in panel a. and b. show that large brokers significantly underperform small brokers, when sorting is done with respect to the first two criteria. In addition, when the market share is measured with respect to the executed volume, $\Delta\Theta$ becomes more negative and more significant. When we move to Panel c., $\Delta\Theta$ remains negative, but insignificant. These results suggest that the conventional belief that the order flow of large brokers conveys information is not necessarily true.

One explanation could be the presence of liquidity suppliers. Nonetheless, since our analysis uses only aggressive trades, we reject this explanation. The presence of noise traders, who, following Black (1986)'s view, mistakenly think they are informed and trade aggressively through large brokers cannot be excluded; however, it is hard to believe that markets leaders trading is dominated by noise traders. We argue that a more plausible explanation is that large brokers attract various types of investors with different levels of sophistication, investment strategies, and ultimately different beliefs. This resembles to a market in which trading is induced by the dispersion of beliefs (see, e.g., Harrison & Kreps (1978) and Harris & Raviv (1993) among others). In this case, although the clientele of large brokers contains informed traders, their sometimes even "orthogonal trading" hides any meaningful signal.

To elaborate on the heterogeneity of investors - and the dispersion of beliefs - argument, we calculate correlated trading statistics. Intuitively, there should be a negative relation between correlated trading and broker clientele diversification; the more heterogeneous the investors are, the more uncorrelated their trading will be. We use the Lakonishok, Shleifer & Vishny (1992) framework[16] to explore this relation. We define the LSV statistic as:

$$H_t(j,i) = |B_t(j,i)/N_t(j,i) - p(j,t)| - AF_t(j,i), \tag{11}$$

where $B(j,i)$ is the number of broker^j trades in stock i during day t that are aggressive purchases, $N(j,i)$ is the number of all trades initiated by broker^j in stock i during day t, $p(j,t)$ is the expected proportion of all broker^j trades that are purchases on day t, and $AF(j,i)$ is an adjustment factor that captures that the first term of the formula can be greater than zero under the null hypothesis of no correlated trading. In our calculation we account for the splitting of large orders effect in the same second, which otherwise will artificially increase the LSV measure. The LSV statistic is computed for each stock-day and then averaged per broker[17].

The results in Panel a. of Table 12, report correlated trading statistics[18], which vary from 2% to 27%, for brokers CDG and RBN, respectively. That means that on average the 52% of the trades initiated by broker CDG every day are on one side of the order book, buys or sells, while this number increases to 77% for RBN. The most interesting result is presented in Panel b: the average size of brokers in the bottom LSV quartile (Q1) is four times the average size of the brokers in the top LSV quartile (Q4). This result supports our argument that large brokers attract very heterogeneous customers. Panel c., shows the underperformance of the portfolio that tracks brokers with very uncorrelated trading.

Overall, our analysis implies that the negative relation between market share and

[16] Recent papers using the same framework are: Grinblatt, Titman & Wermers (1995b), Wermers (1999), Barber, Odean & Zhu (2009a), and Barber, Odean & Zhu (2009b) among others.

[17] To illustrate our approach, suppose that in a given day half of the transactions initiated by broker j are buys and half are sells. We can use this information to infer that broker j clients are trading independently, and the LSV statistic will be close to zero. On the contrary, if 90% percent of broker's j trades are buys, then we would conclude that broker j trading is highly correlated, and the LSV statistic will be greater than zero.

[18] These results are consistent with the previous work of Dorn, Huberman & Sengmueller (2008) and Barber et al. (2009b), who document correlated trading among the clients of a German and a U.S. broker, respectively.

predictive power of brokers' customer order flow, might be explained by the positive relation between market share and heterogeneity of brokers' clients.

4.2 Does Daily Volume *per se* Convey Information?

The above investigation shows that the trading of large brokers, though influential in public opinion's eyes, does not predict future returns. Although the customer order flow of brokers that trade regularly high volumes has on average no predictive power, the information content of daily heavy activity *per se* is an open question. Who initiates volume? Does daily heavy activity contains information on future returns?

To answer these questions, we form a dynamic mean-variance portfolio that uses the average order flow of brokers that trade more heavily on day t in order to predict future returns on day $t+1$, $t+2$, $t+3$ and $t+4$. Table 13 presents the performance against the ANON portfolio and the associated p-values. We find that the customer order flow of brokers who initiate heavy volume on a particular day has very strong predictive power on next day's returns (Panel a). The predictability for returns on day $t+2$ (Panel b), $t+3$ (Panel c) and $t+4$ (Panel d) is statistically insignificant.

We interpret these results as strong evidence that volume acts as a transmission mechanism of new information into prices. In our story, informed investors initiate volume in order to exploit their information advantage. Their activity is then reflected on future prices. However, as our result suggest, the transmission of new information lasts only for one day. Then, prices seem to incorporate all available information. Broadly speaking, there are two dominant views on what generates trading activity. The first one (see e.g., Campbell, Grossman & Wang (1993)), argues that high trading volume is more likely generated by noninformational trading. The second one, argues that it is the activity of informed traders that causes high market volumes. For instance, a recent paper by Martinez & Rosu (2011) shows that a large fraction of trading activity is due to informed trading. Our results favor the latter view.

4.3 Do Investment Style and Sophistication Matter?

In this section, we explore the relation between the trading behavior of investors and the information content of order flow. The underlying idea is the following: If the average customer of brokers exhibits distinct trading behavior, then this may drive brokers' order flow heterogeneity. There are two well-evident investment styles attributed to investors; momentum and contrarian.

We follow the Grinblatt & Keloharju (2000) framework to characterize brokers in terms of their investment style. This framework consists of measuring the difference between the buy ratio of past winning stocks (top quartile) and the buy ratio of past losing stocks (bottom quartile). The buy ratio of broker^j is defined as:

$$\text{Buy Ratio}^j = \frac{\text{Buy Volume}^j}{\text{Buy Volume}^j + \text{Sell Volume}^j}, \tag{12}$$

where all volumes are calculated using the trades initiated by broker^j. We compute daily and hourly buy ratios in order to capture both the daily and intradaily patterns. If the difference is positive (negative), then the broker is viewed as momentum (contrarian) oriented at time t. We calculate buy ratio differences for every time interval, and if the fraction of days (or hours) with positive differences is higher (lower) than 0.50, the broker displays momentum (contrarian) behavior. We analyze statistical significance with both the standard two-sided binomial test and the AR(1)-adjusted binomial test[19] suggested in Grinblatt & Keloharju (2000).

Table 14 presents the fractions of positive buy ratio differences for the intraday and daily horizon, along with the p-values of the associated binomial tests. We observe strong behavioral patterns, both reversal and momentum, at all frequencies. The fraction of positive buy ratio differences varies from 0.34 to 0.60 and 0.28 to 0.68 at the 1 hour and 1 day horizon, respectively. At the daily horizon, FOR, SAB, and UBS are the brokers with the stronger momentum behavior (65%, 68%, and 59%, respectively), and AAL, DBL, and NRD the brokers with the stronger contrarian behavior (39%, 28%, and 35%, respectively). When we move to the intraday frequency, reversal patterns become stronger and the number of significant contrarian brokers doubles.

Our results are broadly consistent with findings of the previous literature. Grinblatt & Keloharju (2000) show that investors in Finland exhibit both contrarian and momentum behavior at the daily horizon, depending on their degree of sophistication, with the least sophisticated investors being contrarian. Linnainmaa (2010) also documents reversal effects using data from HEX. More recently, Heston, Korajczyk & Sadka (2010) find strong intraday return reversals in NYSE that are reversed at the daily frequency, a finding that resembles the weakening of the reversal effects at the daily frequency in our sample.

[19]The z-test statistic of this test is defined as:

$$z = \frac{x - T/2}{\sqrt{T/4 + [(2p-1)^{T+1} - T(2p-1)^2 + (2p-1)(T-1)]/16(1-T)^2]}}, \tag{13}$$

where x is the the fraction of positive buy ratio differences, p is the observed proportion of continuations, and T is number of trading days.

The previous literature[20] suggests that sophisticated market participants exhibit a momentum investment style. Building on this literature, we apply the analysis of the previous section (see Equation 10) and test whether the customer order flow of the statistical significant (p-value$<$ 5%) momentum (daily horizon) brokers contains better information than the one of the statistical significant (p-value$<$ 5%) contrarian (daily horizon) brokers. Table 15 presents the results for Models 1 to 4. In line with the previous evidence, we find that the order flow of the average customer of momentum brokers has statistically strong predictive power for future returns. In contrast, the order flow of contrarian brokers has zero predictive power. As for the performance differences, $\Delta\Theta$, they are positive, varying from 9 to 31 percentage points. Our analysis suggests that the order flow of momentum brokers is more informative than the order flow of contrarian brokers. This finding holds for all models, but it is stronger for Models 2 and 4 and weaker for Models 1 and 3.

The evidence[21] that the order flow of momentum investors conveys information, and the linkage of this finding with the sophistication level is very appealing. Yet we have to prove that this linkage exists in our data. We explore the relation between investment style and sophistication by constructing a measure for the stock picking ability of brokers. We proxy the stock picking ability using, again, the framework of Grinblatt & Keloharju (2000), but this time we examine the buy ratios of future returns. More specifically, if the average buy ratio of future winning stocks (top quartile) exceeds the buy ratio of future losing stocks (bottom quartile), then this provides evidence of high stock picking ability and, thus, high level of sophistication. Future returns are the cumulative daily returns of the next 1 month and 3 months. We compute buy ratio differences for every day and if the fraction of days with positive differences is higher (lower) than 0.50, the broker displays high (low) stock picking ability.

Table 16 presents the fractions of positive buy ratio differences. For both investment horizons, it is evident that the degree of sophistication varies across brokers. The significance of these results varies too. We focus on the two extreme quartiles - the brokers with the highest stock picking ability (Q4) and the brokers with the lowest stock picking ability (Q1) - and compare the average investment style of each group

[20] See e.g., Grinblatt et al. (1995a), Goetzmann & Massa (2002), Griffin, Harris & Topaloglu (2003). Bloomfield et al. (2009) show that short-term momentum is mainly caused by sophisticated informed traders. Grinblatt & Keloharju (2000) find that in HEX the more sophisticated investors are the more momentum is their behavior. Hvidkjaer (2006) reports evidence for informed trading among large traders, whose investment style is momentum.

[21] In unreported results, we examine if the returns of the momentum brokers' portfolio can be explained solely due to a momentum premium. We find that the momentum premium cannot fully explain the momentum brokers' portfolio returns, leaving space for the superiority of their customer order flow information story.

(data taken from Table 14). Table 17 shows that the differences of the two groups are large and statistical significant from zero (p-value$\approx$ 0%): 0.57 vs. 0.48 and 0.56 vs. 0.46 for the 1 month and 3 months horizon, respectively. The results suggest that investors with on average different degrees of sophistication exhibit distinct and opposite investment styles. Therefore, consistent with our expectations and the previous literature, we show that stock picking ability, which proxies the degree of investors' sophistication, is positively related to the degree of momentum trading.

5 Conclusion

The financial crisis of 2008 has triggered a discussion in the finance community about the need to move to more transparent market structures. Policy makers in US and Europe have already taken steps into this direction. One of the dimensions of transparency is the disclosure of counterparty identity. Although early evidence suggests that market anonymity has an positive impact on market liquidity, there are other side-effects on the information content of order flow, the distribution of trading profits, and market efficiency that have not been fully explored empirically.

In this paper, we explore the asset allocation implications of market transparency. Within a simple mean-variance framework and using data from Helsinki Stock Exchange (HEX) we show that investors can benefit remarkably from more transparency. More specifically, we find that a mean-variance investor can improve the allocation of their wealth across the risky assets and the risk-free security by observing the customer order flow of brokers. This is translated into a superior portfolio performance up to 36% (annualized) percentage points for the most parsimonious forecasting model, compared to the benchmark investment scenario, in which investors disregard the broker identity.

A further finding is the quite substantial heterogeneity in the performance of mean-variance portfolios across brokers. Since brokers facilitate trading on behalf of their clients in almost all markets, a large fraction of information is aggregated at the broker level. We examine the determinants of brokers' customer order flow information heterogeneity. We find that the conventional belief that the order flow of large brokers conveys information is not necessarily true. This can be explained by the high dispersion of beliefs among the clients of large brokers. We, also, find that the customer order flow of brokers that trade very actively on a particular day has strong predictive power on next day's returns. Finally, we test whether the information heterogeneity at the broker level is due to their distinct trading behavior. Indeed, we find that brokers that exhibit momentum behavior outperform brokers that exhibit contrarian behavior. This result can be driven by the higher sophistication level of momentum brokers.

References

ANAND, A. & A. SUBRAHMANYAM (2008): "Information and the intermediary: Are market intermediaries informed traders in electronic markets?" *Journal of Financial and Quantitative Analysis*, 43 (01), 1–28.

BARBER, B., T. ODEAN, & N. ZHU (2009a): "Do retail trades move markets?" *Review of Financial Studies*, 22 (1), 151.

—— (2009b): "Systematic noise," *Journal of Financial Markets*, 12 (4), 547–569.

BENVENISTE, L., A. MARCUS, & W. WILHELM (1992): "What's special about the specialist?" *Journal of Financial Economics*, 32 (1), 61–86.

BERGER, D., A. CHABOUD, S. CHERNENKO, E. HOWORKA, & J. WRIGHT (2008): "Order flow and exchange rate dynamics in electronic brokerage system data," *Journal of International Economics*, 75 (1), 93–109.

BLACK, F. (1986): "Noise," *Journal of Finance*, 41 (3), 529–43.

BLOOMFIELD, R., W. TAYLER, & F. ZHOU (2009): "Momentum, reversal, and uninformed traders in laboratory markets," *The Journal of Finance*, 64 (6), 2535–2558.

CAMPBELL, J., S. GROSSMAN, & J. WANG (1993): "Trading volume and serial correlation in stock returns," *The Quarterly Journal of Economics*, 108 (4), 905.

CHORDIA, T., R. ROLL, & A. SUBRAHMANYAM (2002): "Order imbalance, liquidity, and market returns," *Journal of Financial Economics*, 65 (1), 111–130.

COMERTON-FORDE, C. & K. TANG (2009): "Anonymity, liquidity and fragmentation," *Journal of Financial Markets*, 12 (3), 337–367.

COMERTON-FORDE, C., A. FRINO, & V. MOLLICA (2005): "The impact of limit order anonymity on liquidity: Evidence from Paris, Tokyo and Korea," *Journal of Economics and Business*, 57 (6), 528–540.

DELLA CORTE, P., L. SARNO, & G. SESTIERI (2010): "The Predictive Information Content of External Imbalances for Exchange Rate Returns: How Much Is It Worth?" *Review of Economics and Statistics*, forthcoming.

DORN, D., G. HUBERMAN, & P. SENGMUELLER (2008): "Correlated trading and returns," *Journal of Finance*, 63 (2), 885–920.

Easley, D., S. Hvidkjaer, & M. O'Hara (2002): "Is information risk a determinant of asset returns?" *Journal of Finance*, 57 (5), 2185–2221.

Evans, M. & R. Lyons (2002): "Order flow and exchange rate dynamics," *Journal of Political Economy*, 110 (1), 170–180.

—— (2008): "How is macro news transmitted to exchange rates?" *Journal of Financial Economics*, 88 (1), 26–50.

Fishman, M. & F. Longstaff (1992): "Dual trading in futures markets," *Journal of Finance*, 47 (2), 643–671.

Fleming, J., C. Kirby, & B. Ostdiek (2001): "The economic value of volatility timing," *Journal of Finance*, 56 (1), 329–352.

Forster, M. & T. George (1992): "Anonymity in securities markets," *Journal of Financial Intermediation*, 2 (2), 168–206.

Foucault, T., S. Moinas, & E. Theissen (2007): "Does anonymity matter in electronic limit order markets?" *Review of Financial Studies*, 20 (5), 1707.

Friederich, S. & R. Payne (2011): "Post-trade anonymity, order book liquidity and the cost of institutional equity orders," *Working Paper*.

Glosten, L. & P. Milgrom (1985): "Bid, ask and transaction prices in a specialist market with heterogeneously informed traders," *Journal of Financial Economics*, 14 (1), 71–100.

Goetzmann, W. & M. Massa (2002): "Daily momentum and contrarian behavior of index fund investors," *Journal of Financial and Quantitative Analysis*, 37 (03), 375–389.

Goetzmann, W., J. Ingersoll, M. Spiegel, & I. Welch (2007): "Portfolio performance manipulation and manipulation-proof performance measures," *Review of Financial Studies*, 20 (5), 1503–1546.

Goodhart, C. (1988): "The foreign exchange market: a random walk with a dragging anchor," *Economica*, 55 (220), 437–460.

Griffin, J., J. Harris, & S. Topaloglu (2003): "The dynamics of institutional and individual trading," *Journal of Finance*, 58 (6), 2285–2320.

Grinblatt, M. & M. Keloharju (2000): "The investment behavior and performance of various investor types: a study of Finland's unique data set," *Journal of Financial Economics*, 55 (1), 43–67.

Grinblatt, M., S. Titman, & R. Wermers (1995a): "Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior," *The American Economic Review*, 1088–1105.

—— (1995b): "Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior," *The American Economic Review*, 85 (5), 1088–1105.

Harris, M. & A. Raviv (1993): "Differences of opinion make a horse race," *Review of Financial Studies*, 6 (3), 473.

Harrison, J. & D. Kreps (1978): "Speculative investor behavior in a stock market with heterogeneous expectations," *The Quarterly Journal of Economics*, 92 (2), 323.

Hasbrouck, J. (1991): "Measuring the information content of stock trades," *Journal of Finance*, 46 (1), 179–207.

Heston, S., R. Korajczyk, & R. Sadka (2010): "Intraday patterns in the cross-section of stock returns," *Journal of Finance*, 65 (4), 1369–1407.

Hvidkjaer, S. (2006): "A trade-based analysis of momentum," *Review of Financial Studies*, 19 (2), 457–491.

Kyle, A. (1985): "Continuous auctions and insider trading," *Econometrica*, 53 (6), 1315–1335.

Lakonishok, J., A. Shleifer, & R. Vishny (1992): "The impact of institutional trading on stock prices," *Journal of Financial Economics*, 32 (1), 23–43.

Linnainmaa, J. (2010): "Do Limit Orders Alter Inferences about Investor Performance and Behavior?" *Journal of Finance*, 65 (4), 1473–1506.

Linnainmaa, J. & G. Saar (2010): "Lack of Anonymity and the Inference from Order Flow," *CRSP Working Paper*.

Love, R. & R. Payne (2008): "Macroeconomic news, order flows, and exchange rates," *Journal of Financial and Quantitative Analysis*, 43 (2), 467–488.

Lyons, R. (1997): "A simultaneous trade model of the foreign exchange hot potato," *Journal of International Economics*, 42 (3-4), 275–298.

Marquering, W. & M. Verbeek (2004): "The economic value of predicting stock index returns and volatility," *Journal of Financial and Quantitative Analysis*, 39 (2), 407–429.

Martinez, V. & I. Rosu (2011): "High-Frequency Traders, News and Volatility," *Working Paper, SSRN.*

Nolte, I. & S. Nolte (2010): "How do individual investors trade?" *European Journal of Finance (forthcoming).*

Pagano, M. & A. Roell (1993): "Front-running and stock market liquidity," *Financial markets' liberalisation and the role of banks*, 178.

Pasquariello, P. & C. Vega (2007): "Informed and strategic order flow in the bond markets," *Review of Financial Studies*, 20 (6), 1975–2019.

Payne, R. (2003): "Information transmission in inter-dealer foreign exchange transactions," *Journal of International Economics*, 61 (2), 307–329.

Peiers, B. (1997): "Informed traders, intervention, and price leadership: A deeper view of the microstructure of the foreign exchange market," *Journal of Finance*, 52 (4), 1589–1614.

Rime, D., L. Sarno, & E. Sojli (2010): "Exchange rate forecasting, order flow and macroeconomic information," *Journal of International Economics*, 80 (1), 72–88.

Rindi, B. (2008): "Informed traders as liquidity providers: Anonymity, liquidity and price formation," *Review of Finance*, 12 (3), 497.

Roell, A. (1990): "Dual-capacity trading and the quality of the market," *Journal of Financial Intermediation*, 1 (2), 105–124.

Sapp, S. (2002): "Price leadership in the spot foreign exchange market," *Journal of Financial and Quantitative Analysis*, 37 (3), 425–448.

Wermers, R. (1999): "Mutual fund herding and the impact on stock prices," *Journal of Finance*, 54 (2), 581–622.

6 Appendix

Table 1: List of Brokers in HEX

These are the brokers of the HEX, ranked alphabetically. The first 2 columns are the full name and the code of each broker, respectively. The third column is the nationality of each broker. The letter R in the last column identifies the remote members. (Source: `http://nordic.nasdaqomxtrader.com/membershipservices/membershiplist`)

Broker Name	*Code*	*Country*	*RM*
Ålandsbanken Abp	AAL	Finland	
ABN AMRO Clearing Bank N.V	FOR	Netherlands	R
BNP Paribas Arbitrage SNC	BPP	France	R
Carnegie Investment Bank AB	CAR	Sweden	
Citadel Securities (Europe) Limited	CDG	UK	R
Citigroup Global Markets Limited	SAB	UK	R
Crédit Agricole Cheuvreux Nordic AB	CDV	Sweden	
Credit Suisse Securities (Europe) Ltd	CSB	UK	R
Danske Bank A/S	DDB	Denmark	
Deutsche Bank AG	DBL	UK	R
Evli Bank Abp	EVL	Finland	
FIM Bank Ltd.	FIM	Finland	
Goldman Sachs International	GSI	UK	R
Instinet Europe Limited	INT	UK	R
JP Morgan Securities Ltd	JPM	UK	R
Knight Capital Europe Limited	KEM	UK	R
Merrill Lynch International	MLI	UK	R
Morgan Stanley Co. International Ltd.	MSI	UK	R
NeoNet Securities AB	NEO	Sweden	
Nomura International plc	NIP	UK	R
Nordea Bank Finland Plc	NRD	Finland	
Nordnet Bank AB	NON	Sweden	
Pohjola Bank Plc	OPS	Finland	
SAXO-E*TRADE Bank A/S	DIF	Denmark	
Skandinaviska Enskilda Banken AB	ENS	Sweden	
Société Générale S.A.	SGP	France	R
Swedbank AB	SWB	Sweden	
Svenska Handelsbanken AB	SHB	Sweden	
The Royal Bank of Scotland N.V.	RBN	UK	R
UBS Limited	UBS	UK	R
UB Securities Limited	UB	Finland	

Table 2: Summary Statistics.

The table reports the name, mean (%), standard deviation (%), maximum (%), minimum (%), skewness, and kurtosis of the returns of the 15 most liquid stocks of HEX25 index. All these statistics are calculated over the daily interval. The first set of statistics corresponds to the period 03/30/2010 - 02/28/2011; the next two sets of statistics correspond to the two subperiods of 03/30/2010 - 08/09/2010 and 08/10/2010 - 02/28/2011, respectively. The last column, reports the total aggressive turnover ('000,000) traded in each stock during the whole sample period. We choose to perform our analysis only on the 15 most liquid stocks (with respect to the turnover) in order to circumvent problems arising from the low number of transactions of some brokers.

		03/30/2010 - 02/28/2011						03/30/2010 - 08/09/2010						08/10/2010 - 02/28/2011						
#	*Name*	*Mean*	*Std*	*Max*	*Min*	*Skew*	*Kurt*	*Mean*	*Std*	*Max*	*Min*	*Skew*	*Kurt*	*Mean*	*Std*	*Max*	*Min*	*Skew*	*Kurt*	*Turn*
1	***Elisa***	0.03	1.36	7.08	-6.23	0.40	8.19	0.00	1.51	7.08	-3.65	1.05	7.39	0.05	1.26	5.00	-6.23	-0.32	8.58	1,624
2	***Fortum***	0.09	1.34	4.75	-5.60	-0.28	4.81	0.02	1.61	4.75	-5.60	-0.13	4.28	0.13	1.13	2.63	-4.38	-0.39	4.09	7,059
3	***Kone***	0.11	1.57	5.77	-4.35	0.20	4.08	0.17	1.85	5.77	-4.25	0.24	3.75	0.07	1.36	4.03	-4.35	0.04	3.47	3,489
4	***Konecranes***	0.16	2.05	8.74	-4.69	0.78	5.23	0.16	2.24	8.53	-4.62	0.64	4.60	0.16	1.93	8.74	-4.69	0.90	5.61	1,641
5	***Metso***	0.18	2.49	9.35	-6.68	0.14	3.76	0.26	3.14	9.35	-6.68	0.15	2.96	0.12	1.97	5.75	-4.89	-0.06	3.34	5,090
6	***Neste Oil***	-0.01	1.71	6.77	-6.65	-0.23	4.66	-0.13	2.06	6.77	-6.02	0.06	3.66	0.07	1.44	3.23	-6.65	-0.59	5.40	2,318
7	***Nokia***	-0.26	2.36	6.09	-15.33	-2.45	16.71	-0.50	2.57	5.25	-15.30	-2.60	14.78	-0.10	2.21	6.09	-15.33	-2.20	17.71	38,370
8	***Nokian Renkaat***	0.18	2.06	9.88	-5.89	0.58	5.07	0.23	2.42	9.88	-4.47	0.81	4.88	0.15	1.80	5.89	-5.89	0.11	3.87	2,990
9	***Outokumpu***	-0.10	2.15	8.44	-6.63	0.23	4.26	-0.16	2.67	8.44	-6.63	0.28	3.53	-0.06	1.75	5.70	-5.57	0.16	3.76	3,374
10	***Outotec***	0.18	2.51	11.40	-8.14	0.28	4.59	0.12	3.00	11.40	-8.14	0.31	4.49	0.22	2.14	6.09	-4.89	0.24	3.04	2,616
11	***Rautarukki K***	0.01	2.15	8.60	-5.85	0.49	4.39	-0.06	2.56	8.60	-5.85	0.67	4.24	0.06	1.85	5.43	-4.36	0.19	3.34	1,978
12	***Sampo A***	0.06	1.53	8.89	-6.01	0.51	8.12	-0.01	2.01	8.89	-6.01	0.54	6.49	0.10	1.12	3.51	-2.52	0.38	3.23	5,267
13	***Stora Enso R***	0.16	2.23	8.07	-6.68	0.12	3.86	0.19	2.55	7.92	-6.68	-0.18	3.50	0.14	2.00	8.07	-3.85	0.49	3.81	6,095
14	***UPM-Kymmene***	0.16	2.04	8.44	-5.76	0.07	4.27	0.19	2.22	8.44	-5.62	-0.03	4.64	0.15	1.92	6.56	-5.76	0.16	3.58	6,512
15	***Wärtsilä Abp***	0.17	2.08	9.61	-5.24	0.32	4.78	0.15	2.45	9.61	-5.08	0.57	4.48	0.18	1.81	6.38	-5.24	-0.08	4.06	3,280

Table 3: A 5 Second Slice of the Transaction Data of NOK1V

This is a 5 second slice of the transaction data of NOK1V. The first two columns are the date and time expressed as month/day/year and hour:minute:second, respectively. The third column is the type of the transaction, which can be Best Bid, Best Ask or Trade. The next two columns are the price (in euros) and the size of the transaction. The last two columns are the Broker Buy Code and the Broker Sell Code.

Date	*Time*	*Type*	*Price*	*Size*	*Broker Buy*	*Broker Sell*
4012010	08:03:51	BEST BID	11.6	15,531		
4012010	08:03:51	BEST BID	11.6	13,531		
4012010	08:03:53	BEST ASK	11.61	14,876		
4012010	08:03:55	BEST ASK	11.61	9,876		
4012010	08:03:55	BEST ASK	11.61	6,876		
4012010	08:03:55	BEST BID	11.6	12,331		
4012010	08:03:55	TRADE	11.61	1,161	ENS	ENS
4012010	08:03:55	TRADE	11.61	39	ENS	NON
4012010	08:03:55	BEST ASK	11.61	5,676		
4012010	08:03:56	BEST BID	11.6	12,305		
4012010	08:03:56	TRADE	11.6	26	ENS	NON
4012010	08:03:56	TRADE	11.6	1,305	ENS	NON
4012010	08:03:56	TRADE	11.6	669	ENS	NON

Table 4: Statistical Performance of The First (M1) Forecasting Model.

The table gives an indication of the statistical performance of the first order flow forecasting model (M1) by reporting the heteroskedasticity and autocorrelation corrected (Newey-West) t-statistics of the order flow measure (OF) coefficient in the initial estimation window (03/30/2010–08/09/2010, 86 days), and across the 15 most liquid stocks of HEX25 index. OF is defined as the daily difference between the buyer-initiated and seller-initiated volume. In bold are the statistical significant (p-value< 10%) t-statistics.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]
AAL	1.02	**5.44**	-0.50	-0.46	1.04	**2.32**	-1.55	0.38	**3.01**	**4.68**	0.74	1.09	-0.40	**3.80**	**-1.93**
BPP	0.89	**-2.87**	-1.45	-0.26	-1.29	-0.61	-1.38	-0.53	**2.03**	-0.66	0.45	-1.55	-0.77	0.51	-0.53
CAR	-0.09	**1.67**	0.24	**-1.99**	1.53	-0.35	-0.60	**2.76**	1.31	0.28	1.57	0.30	-0.17	**1.83**	**6.42**
CDG	0.95	0.15	0.84	**-2.58**	-0.12	-0.52	-0.33	**-1.65**	0.21	-0.36	-1.52	-0.08	0.04	-0.67	-1.50
CDV	0.22	0.13	**-1.72**	-0.58	-1.34	-0.05	0.80	0.92	0.84	-1.38	-0.67	0.34	-0.11	1.26	-1.40
CSB	-0.89	-0.59	-0.93	0.65	-0.75	-0.91	-0.73	-1.51	-1.29	-0.63	0.07	1.27	-0.10	**1.77**	-0.51
DBL	0.17	0.24	-0.68	1.49	-0.25	-0.77	**2.37**	**-1.86**	0.81	0.10	0.21	0.89	-0.10	1.07	**1.74**
DDB	-0.30	**2.69**	**-1.84**	1.19	**2.04**	-1.08	-1.44	1.20	-0.59	0.91	0.42	-0.55	-0.38	-0.08	-0.29
DIF	0.01	-0.04	**-1.66**	-1.40	-0.32	-0.15	-0.16	**-1.84**	**2.84**	0.96	1.15	**-1.67**	1.07	-0.49	**-3.69**
ENS	0.65	-0.43	-1.63	**1.70**	0.32	-0.02	1.02	0.12	-0.09	1.02	-0.71	**2.10**	-0.68	1.29	1.04
EVL	0.84	0.23	-1.62	-0.26	**1.93**	0.06	**2.11**	1.42	0.58	0.23	-0.89	-0.20	1.36	0.31	0.88
FIM	-0.33	-0.50	**-3.09**	1.06	**-1.91**	-0.44	-0.11	0.15	-0.13	0.17	-0.08	**-1.90**	0.04	-0.68	1.16
FOR	0.30	0.20	0.42	-0.28	-0.75	-0.93	**-1.71**	-0.59	-0.52	-1.42	1.04	**4.73**	**1.90**	0.93	0.77
GSI	1.52	0.31	0.38	0.46	**2.78**	-1.22	-1.07	-1.31	**2.03**	**-1.97**	-0.90	1.56	1.16	-0.57	-0.44
INT	-0.18	**-1.94**	-1.06	**12.84**	-0.26	-1.52	-0.51	0.55	-0.80	-0.77	**-4.75**	-0.56	0.01	0.02	-0.22
KEM	**-1.95**	**2.94**	-0.51	-0.68	1.25	0.34	0.57	-1.20	-0.28	**3.51**	-0.28	0.34	-0.40	-0.42	-0.04
JPM	1.11	**4.63**	-0.05	**-1.84**	0.11	-0.49	**2.21**	-1.02	-0.33	**-2.00**	-0.22	0.12	-0.04	0.56	-0.46
MLI	**1.85**	-0.45	-0.76	0.52	0.93	0.13	-0.11	-0.37	1.02	0.45	-0.86	-0.42	1.34	-1.50	-0.56
MSI	-1.34	0.90	-1.27	-0.52	-0.12	-0.38	-0.14	-1.20	-0.31	0.48	-1.58	0.92	-0.98	-1.34	0.77
NEO	-0.43	-0.57	1.15	**2.42**	**-3.25**	0.51	-0.73	-0.01	0.28	**-1.91**	-1.64	**2.19**	1.28	0.21	**-2.31**
NIP	0.86	0.94	0.56	-1.28	0.34	0.28	0.98	1.62	-0.71	-1.58	0.72	-0.63	1.03	0.99	-0.90
NON	-1.12	0.24	1.25	-0.23	0.92	1.63	-0.22	-0.21	-0.12	-0.64	0.27	**-3.58**	0.40	0.39	-0.23
NRD	1.18	0.13	0.51	-1.40	0.83	**-1.81**	0.36	-1.21	-0.16	-1.17	-0.86	0.89	0.10	-1.42	-0.27
OPS	-0.31	0.33	0.41	-0.06	0.04	0.50	0.50	-0.60	-1.09	-0.54	-0.33	-0.93	-0.60	0.24	-0.48
RBN	-0.23	**-3.29**	0.24	1.04	0.64	-0.62	0.05	-0.52	**1.91**	1.27	-0.98	-0.28	-1.42	-0.79	**1.84**
SAB	1.10	0.09	0.58	**-1.96**	-1.32	-1.02	-0.14	**-2.18**	1.02	-0.10	-1.30	0.79	0.40	1.00	1.37
SGP	-0.89	0.15	0.21	-0.45	0.21	0.10	0.09	0.21	1.52	-0.72	1.39	0.81	**-1.82**	0.80	-0.55
SHB	0.67	0.33	0.03	0.71	1.23	**2.67**	**2.25**	**1.95**	-0.45	0.52	**1.72**	1.60	**-2.50**	**2.52**	0.39
SWB	**-5.63**	**2.71**	0.41	**3.19**	**1.83**	-0.59	0.11	**-5.90**	0.32	1.46	0.15	-0.91	-0.60	1.62	0.32
UB	-0.51	0.18	**1.88**	0.49	-0.77	-1.10	-0.74	**4.22**	-0.17	**3.12**	1.30	1.24	**3.26**	0.89	1.08
UBS	0.87	-0.44	-1.12	0.33	-0.04	-0.77	-0.16	-0.73	**-2.70**	-0.31	0.73	0.06	0.88	1.46	-0.05

Table 5: Performance in the Initial Period.

The table presents the performance of the mean variance portfolios in the initial period: 03/30/2010–08/09/2010 (86 days). There is one portfolio for each broker. The ANON portfolio is the one that disregards the broker identity, and it is the benchmark portfolio. The investment scenario is based on a risk-averse investor, who maximizes his expected portfolio return subject to an annual target volatility $\sigma_p = 10\%$. Every day, the investor forecasts next day's returns using the order flow models $M1 - M4$ described in section 2, and then rebalances his portfolio weights. The order flow models are estimated once using all 86 days. We present: the annualized Sharpe ratio (SR) of each portfolio and Θ. Θ is the difference between the broker^j's and the ANON's performance measure (MPPM) of Goetzmann et al. (2007), which is expressed in (annualized) percentage points and is for $\gamma = 6$. It can be viewed as the maximum performance fee an investor is willing to pay to switch from the ANON portfolio to the broker^j's portfolio. When $\Theta > 0$, market transparency yields positive economic value to mean variance investors. Following Goetzmann et al. (2007), we test whether the broker^j's portfolio significantly outperforms the ANON portfolio and report the p-values in square brackets.

	M1			M2			M3			M4		
Broker	*SR*	Θ	*p-val*	*SR*	Θ	*p-val*	*SR*	Θ	*p-val*	*SR*	Θ	*p-val*
AAL	4.82	-1	[0.53]	5.57	2	[0.46]	6.12	5	[0.39]	6.99	8	[0.32]
BPP	5.24	11	[0.32]	6.07	12	[0.29]	5.96	13	[0.28]	7.05	18	[0.17]
CAR	4.66	-8	[0.66]	5.08	-5	[0.61]	6.21	1	[0.48]	6.32	0	[0.49]
CDG	6.23	13	[0.26]	8.39	32	[0.07]	7.48	15	[0.19]	9.72	37	[0.03]
CDV	3.46	-16	[0.83]	5.55	2	[0.45]	5.06	-6	[0.64]	7.07	9	[0.26]
CSB	4.75	1	[0.48]	4.58	2	[0.46]	6.39	10	[0.30]	5.41	4	[0.43]
DBL	7.98	28	[0.06]	8.43	31	[0.05]	8.56	28	[0.05]	9.50	31	[0.04]
DDB	7.43	34	[0.06]	7.67	40	[0.04]	8.37	35	[0.04]	8.71	40	[0.02]
DIF	5.44	-8	[0.67]	6.47	3	[0.43]	7.16	4	[0.40]	7.72	11	[0.27]
ENS	5.66	5	[0.40]	6.51	19	[0.23]	7.22	12	[0.28]	7.65	23	[0.16]
EVL	4.04	-22	[0.91]	4.35	-16	[0.80]	5.54	-14	[0.82]	5.20	-16	[0.82]
FIM	4.42	-7	[0.64]	4.87	-7	[0.65]	5.66	0	[0.51]	5.88	-2	[0.53]
FOR	7.62	24	[0.12]	8.61	20	[0.12]	8.99	26	[0.07]	9.72	22	[0.07]
GSI	3.89	-17	[0.83]	5.62	-1	[0.53]	5.72	-8	[0.70]	7.09	2	[0.45]
INT	5.99	8	[0.35]	5.48	0	[0.50]	7.59	18	[0.15]	6.65	2	[0.46]
JPM	5.25	11	[0.27]	7.22	29	[0.09]	6.21	14	[0.22]	8.03	31	[0.07]
KEM	5.30	-1	[0.52]	5.85	-1	[0.52]	7.13	8	[0.30]	7.76	8	[0.28]
MLI	5.65	-1	[0.51]	5.17	-6	[0.62]	6.76	5	[0.39]	6.66	1	[0.48]
MSI	6.82	18	[0.15]	6.13	5	[0.40]	7.87	21	[0.08]	6.98	5	[0.39]
NEO	5.33	-4	[0.58]	6.16	5	[0.40]	6.79	6	[0.38]	7.19	12	[0.27]
NIP	4.96	-2	[0.53]	6.25	20	[0.19]	7.03	12	[0.26]	7.56	28	[0.10]
NON	5.67	14	[0.26]	6.94	29	[0.11]	7.42	27	[0.10]	8.04	30	[0.08]
NRD	6.86	23	[0.11]	8.32	30	[0.06]	7.73	26	[0.08]	8.76	31	[0.04]
OPS	3.06	-20	[0.79]	4.05	-13	[0.72]	6.10	1	[0.48]	5.94	-2	[0.54]
RBN	5.59	8	[0.36]	6.25	16	[0.23]	6.63	13	[0.26]	7.28	22	[0.15]
SAB	5.94	8	[0.35]	6.79	13	[0.26]	6.93	11	[0.28]	7.27	15	[0.22]
SGP	4.46	-11	[0.73]	4.95	-6	[0.63]	6.78	1	[0.49]	7.22	5	[0.37]
SHB	6.91	22	[0.16]	5.82	10	[0.34]	8.16	25	[0.10]	7.57	14	[0.25]
SWB	5.78	15	[0.21]	7.84	41	[0.03]	7.14	22	[0.09]	8.73	43	[0.01]
UB	5.87	11	[0.30]	6.59	14	[0.26]	7.03	13	[0.24]	7.72	16	[0.20]
UBS	5.55	3	[0.44]	7.41	16	[0.17]	6.93	14	[0.20]	8.68	22	[0.06]
ANON	5.42	0	[0.50]	5.42	0	[0.50]	7.13	0	[0.50]	7.13	0	[0.50]

Table 6: Does Broker Identity Convey Information?.

The table presents the performance of the mean-variance portfolios, using a recursive (out-of-sample) regression estimation, which is based on a window of expanding size. The period is 08/10/2010–02/28/2011 (124 days). There is one portfolio for each broker. The ANON portfolio is the one that disregards the broker identity, and it is the benchmark portfolio. The investment scenario is based on a risk-averse investor, who maximizes his expected portfolio return subject to an annual target volatility $\sigma_p = 10\%$. Every day, the investor forecasts next day's returns using the order flow models $M1 - M4$ described in section 2, and then rebalances his portfolio weights. We present: the annualized Sharpe ratio (SR) of each portfolio and Θ. Θ is the difference between the broker^j's and the ANON's performance measure (MPPM) of Goetzmann et al. (2007), which is expressed in (annualized) percentage points and is for $\gamma = 6$. It can be viewed as the maximum performance fee an investor is willing to pay to switch from the ANON portfolio to the broker^j's portfolio. When $\Theta > 0$, market transparency yields positive economic value to mean variance investors. Following Goetzmann et al. (2007), we test whether the broker^j's portfolio significantly outperforms the one that disregards the broker identity and report the p-values in square brackets.

	M1			M2			M3			M4		
Broker	*SR*	Θ	*p-val*	*SR*	Θ	*p-val*	*SR*	Θ	*p-val*	*SR*	Θ	*p-val*
AAL	0.37	10	[0.29]	-0.05	6	[0.38]	-1.10	4	[0.42]	-1.24	2	[0.46]
BPP	-0.47	2	[0.44]	0.00	7	[0.33]	-1.25	3	[0.43]	-0.38	12	[0.22]
CAR	2.93	34	[0.01]	0.72	14	[0.20]	1.62	30	[0.01]	0.29	18	[0.13]
CDG	1.41	23	[0.14]	1.66	22	[0.11]	0.75	23	[0.13]	0.61	21	[0.11]
CDV	1.45	23	[0.08]	0.94	17	[0.17]	0.75	23	[0.07]	0.29	17	[0.16]
CSB	0.51	12	[0.23]	1.45	23	[0.09]	-0.51	6	[0.34]	0.68	22	[0.10]
DBL	-0.85	-4	[0.59]	-1.11	-8	[0.67]	-1.63	-5	[0.62]	-1.52	-4	[0.59]
DDB	0.18	8	[0.31]	0.72	14	[0.20]	-0.03	14	[0.19]	0.40	19	[0.13]
DIF	-0.55	2	[0.45]	-0.09	6	[0.36]	-1.00	6	[0.35]	-0.56	11	[0.27]
ENS	0.81	15	[0.12]	-0.79	-1	[0.53]	0.40	18	[0.07]	-1.15	3	[0.43]
EVL	1.29	20	[0.15]	0.88	16	[0.19]	0.39	19	[0.16]	0.35	18	[0.15]
FIM	0.17	8	[0.34]	0.47	11	[0.27]	-0.48	9	[0.31]	-0.26	13	[0.24]
FOR	2.00	27	[0.04]	0.71	14	[0.17]	1.03	26	[0.05]	-0.01	15	[0.14]
GSI	1.33	20	[0.09]	1.01	17	[0.15]	0.71	22	[0.07]	0.74	22	[0.08]
INT	1.03	18	[0.13]	0.97	17	[0.14]	0.34	18	[0.12]	0.38	18	[0.12]
JPM	2.21	33	[0.04]	0.65	13	[0.24]	1.53	33	[0.04]	0.01	15	[0.20]
KEM	1.23	19	[0.08]	0.75	14	[0.18]	0.26	18	[0.09]	-0.44	11	[0.23]
MLI	0.10	8	[0.32]	0.01	7	[0.33]	-0.88	7	[0.34]	-0.49	11	[0.25]
MSI	0.75	14	[0.21]	0.60	13	[0.27]	0.19	17	[0.17]	0.13	16	[0.21]
NEO	0.63	13	[0.17]	0.70	14	[0.20]	0.04	16	[0.12]	0.00	15	[0.17]
NIP	-0.22	3	[0.42]	-0.49	0	[0.49]	-1.02	2	[0.45]	-0.80	6	[0.36]
NON	-1.61	-10	[0.73]	-2.04	-15	[0.80]	-2.03	-7	[0.68]	-2.61	-13	[0.78]
NRD	2.44	29	[0.02]	2.79	31	[0.03]	1.53	29	[0.02]	1.28	27	[0.06]
OPS	0.30	9	[0.29]	0.74	14	[0.22]	-0.06	13	[0.20]	0.36	18	[0.15]
RBN	2.64	36	[0.03]	2.90	38	[0.03]	2.28	38	[0.02]	2.68	43	[0.02]
SAB	0.81	15	[0.22]	0.90	16	[0.19]	-0.11	13	[0.24]	-0.11	14	[0.22]
SGP	0.19	8	[0.31]	2.14	28	[0.07]	-0.55	8	[0.30]	1.24	27	[0.07]
SHB	2.23	29	[0.03]	1.61	23	[0.09]	1.45	30	[0.02]	1.08	26	[0.05]
SWB	2.50	36	[0.03]	1.93	29	[0.06]	1.91	36	[0.02]	1.21	28	[0.07]
UB	0.16	8	[0.29]	0.54	12	[0.23]	-0.15	14	[0.17]	0.34	18	[0.12]
UBS	1.42	22	[0.10]	2.08	29	[0.08]	0.52	20	[0.12]	1.07	26	[0.10]
ANON	-0.57	0	[0.50]	-0.57	0	[0.50]	-1.39	0	[0.50]	-1.39	0	[0.50]

Table 7: Correlations between Brokers' Order Flow and the ANON Order Flow.

The table reports the correlations between the brokers' order flow and the aggregate market order flow across the 15 most liquid stocks (first row) of HEX25 index. ANON order flow corresponds to the aggregate market order flow. Order flow is measured as the daily difference between the number of buyer-initiated and seller-initiated volume. We construct order flow measures of brokers (first column) over daily intervals in the period 03/30/2010 - 02/28/2011.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]
AAL	-0.32	0.15	-0.11	0.02	0.08	-0.01	0.12	0.11	-0.14	-0.08	0.10	-0.07	0.10	-0.03	0.10
BPP	0.33	0.11	0.22	0.16	0.18	0.20	0.42	0.05	0.13	0.05	0.20	0.10	0.39	0.19	0.01
CAR	0.04	0.06	0.12	0.06	0.22	-0.01	0.02	0.31	0.01	0.35	0.08	-0.09	-0.04	-0.10	0.29
CDG	0.07	0.16	0.35	0.29	0.29	0.18	0.33	0.37	0.28	0.46	0.36	0.30	0.26	0.35	0.22
CDV	0.04	0.06	0.36	0.19	0.14	0.08	0.07	0.00	0.11	0.12	0.07	-0.03	-0.02	0.13	0.06
CSB	0.22	-0.04	0.45	0.37	0.23	0.17	0.34	0.32	0.34	0.33	0.34	0.30	0.16	0.20	0.09
DBL	0.39	0.15	0.22	0.22	0.18	0.40	-0.04	0.35	0.28	0.19	0.14	0.11	-0.06	0.29	0.12
DDB	-0.08	0.09	0.02	0.12	0.07	0.35	0.01	0.00	0.00	0.08	0.27	0.05	0.29	0.28	0.36
DIF	-0.15	0.03	0.14	0.09	0.11	0.26	0.41	0.07	0.15	-0.02	-0.02	-0.05	0.10	0.23	0.00
ENS	0.82	0.22	0.37	0.37	0.48	0.08	0.19	0.17	0.28	0.18	0.09	0.16	0.36	0.33	0.27
EVL	0.11	-0.02	0.04	0.06	0.26	0.07	0.06	0.13	0.13	0.08	0.12	0.02	0.28	0.03	0.14
FIM	-0.03	0.32	-0.06	0.23	0.16	0.26	0.04	0.22	0.22	0.09	0.24	0.09	0.15	-0.02	0.20
FOR	0.32	0.32	0.22	0.14	0.21	0.35	0.54	0.20	0.42	0.25	0.34	0.53	0.22	0.42	0.18
GSI	0.43	0.37	0.08	0.14	0.17	0.29	0.19	0.27	0.09	0.14	0.29	0.32	0.14	0.31	0.17
INT	0.04	0.03	0.24	0.04	0.14	0.10	0.10	0.34	-0.03	0.24	0.14	0.03	0.16	0.12	-0.11
JPM	0.01	0.30	0.29	0.08	0.13	0.12	0.06	0.34	-0.07	0.24	0.20	0.13	0.02	0.16	0.09
KEM	0.02	0.11	-0.05	0.05	0.16	-0.07	0.09	-0.04	0.29	0.05	0.16	0.04	-0.08	0.03	0.19
MLI	0.22	0.16	0.20	0.34	0.34	-0.02	0.29	0.29	0.12	0.20	0.14	0.16	-0.09	0.19	-0.02
MSI	0.46	0.26	0.29	0.15	0.08	0.10	0.12	0.31	0.39	0.19	0.32	0.26	0.15	0.16	0.11
NEO	0.00	0.12	-0.06	0.07	0.04	0.02	-0.12	0.00	0.09	0.14	0.06	0.07	-0.05	-0.02	0.07
NIP	0.05	0.28	0.23	0.22	0.14	0.19	0.04	0.37	0.19	0.28	0.24	0.41	0.21	0.32	0.36
NON	-0.47	0.02	-0.12	0.05	0.00	0.10	0.00	0.15	0.00	0.07	-0.01	-0.09	0.17	0.15	-0.04
NRD	-0.04	0.04	0.02	0.19	0.16	0.19	0.22	0.04	0.00	0.05	0.21	0.12	0.20	0.07	0.14
OPS	0.05	0.13	0.14	0.26	-0.07	0.13	-0.01	0.10	-0.02	-0.03	0.23	0.13	0.15	0.07	0.03
RBN	0.15	-0.10	0.03	0.01	0.13	0.04	0.03	0.08	0.20	0.01	0.08	0.11	0.01	-0.03	0.00
SAB	0.16	0.35	0.50	0.38	0.41	0.30	0.19	0.32	0.34	0.44	0.43	0.11	0.11	0.38	0.09
SGP	0.09	0.18	-0.01	0.03	-0.05	-0.05	0.25	0.04	0.08	0.08	0.15	0.32	0.25	0.09	0.09
SHB	0.62	0.22	0.13	0.09	0.09	0.14	-0.01	0.27	0.26	0.39	0.14	0.20	0.17	0.11	0.09
SWB	0.18	0.05	0.15	0.06	0.08	0.06	0.09	0.21	-0.01	0.07	0.00	-0.04	0.11	0.14	0.05
UB	-0.28	0.07	0.19	0.08	-0.18	0.13	0.22	-0.01	0.02	0.04	0.03	0.01	0.18	-0.05	-0.11
UBS	-0.59	0.11	0.13	0.15	0.31	0.24	0.40	0.31	0.15	0.28	0.10	0.19	-0.09	0.20	0.10

Table 8: The Impact of Transactions Costs - Initial Period.

The table presents the impact of transaction costs on the performance of the mean-variance portfolios in the initial period: 03/30/2010–08/09/2010 (86 days). There is one portfolio for each broker, while the portfolio under anonymity (ANON) is the benchmark portfolio. Same as before, the investment scenario is based on an annual target volatility $\sigma_p = 10\%$, and daily portfolio weights rebalancing, using the order flow models $M1 - M4$ described in section 2. The order flow models are estimated once using all 86 days. We present the annualized Θ for 3 levels of transaction costs (TC); Θ_0 for $TC = 0bps$, Θ_{10} for $TC = 10bps$, and Θ_{30} for $TC = 30bps$. Θ is the difference between the broker^j's and the anonymous market's performance measure (MPPM) of Goetzmann et al. (2007), which is expressed in percentage points and is for $\gamma = 6$. Following Goetzmann et al. (2007), we test whether the broker^j's portfolio significantly outperforms the one that disregards broker identities and report the p-values in square brackets.

	M1						**M2**						**M3**						**M4**					
	Θ_0	*p-val*	Θ_{10}	*p-val*	Θ_{30}	*p-val*	Θ_0	*p-val*	Θ_{10}	*p-val*	Θ_{30}	*p-val*	Θ_0	*p-val*	Θ_{10}	*p-val*	Θ_{30}	*p-val*	Θ_0	*p-val*	Θ_{10}	*p-val*	Θ_{30}	*p-val*
AAL	-1	[0.53]	10	[0.30]	33	[0.05]	2	[0.46]	4	[0.43]	7	[0.35]	5	[0.39]	13	[0.23]	29	[0.05]	8	[0.32]	8	[0.32]	8	[0.32]
BPP	11	[0.32]	16	[0.25]	25	[0.14]	12	[0.29]	9	[0.33]	4	[0.42]	13	[0.28]	15	[0.24]	21	[0.17]	18	[0.17]	15	[0.22]	8	[0.33]
CAR	-8	[0.66]	0	[0.51]	14	[0.22]	-5	[0.61]	-3	[0.57]	1	[0.47]	1	[0.48]	7	[0.34]	19	[0.13]	0	[0.49]	1	[0.46]	4	[0.41]
CDG	13	[0.26]	18	[0.18]	30	[0.07]	32	[0.07]	30	[0.08]	26	[0.12]	15	[0.19]	20	[0.12]	30	[0.04]	37	[0.03]	36	[0.03]	33	[0.04]
CDV	-16	[0.83]	-1	[0.53]	29	[0.05]	2	[0.45]	9	[0.29]	22	[0.07]	-6	[0.64]	5	[0.38]	27	[0.05]	9	[0.26]	14	[0.15]	24	[0.03]
CSB	1	[0.48]	7	[0.36]	19	[0.17]	2	[0.46]	0	[0.49]	-3	[0.56]	10	[0.30]	14	[0.22]	23	[0.11]	4	[0.43]	1	[0.47]	-3	[0.56]
DBL	28	[0.06]	38	[0.02]	59	[0.00]	31	[0.05]	32	[0.04]	36	[0.03]	28	[0.05]	36	[0.02]	51	[0.00]	31	[0.04]	32	[0.04]	35	[0.03]
DDB	34	[0.06]	41	[0.03]	55	[0.01]	40	[0.04]	38	[0.05]	34	[0.07]	35	[0.04]	41	[0.02]	52	[0.01]	40	[0.02]	38	[0.03]	36	[0.04]
DIF	-8	[0.67]	-1	[0.52]	14	[0.23]	3	[0.43]	4	[0.41]	7	[0.37]	4	[0.40]	8	[0.32]	15	[0.19]	11	[0.27]	10	[0.27]	9	[0.30]
ENS	5	[0.40]	13	[0.26]	28	[0.08]	19	[0.23]	19	[0.23]	19	[0.23]	12	[0.28]	15	[0.22]	22	[0.13]	23	[0.16]	22	[0.17]	20	[0.20]
EVL	-22	[0.91]	-8	[0.69]	20	[0.11]	-16	[0.80]	-15	[0.78]	-12	[0.74]	-14	[0.82]	-4	[0.61]	15	[0.16]	-16	[0.82]	-15	[0.81]	-14	[0.79]
FIM	-7	[0.64]	1	[0.47]	18	[0.19]	-7	[0.65]	-5	[0.60]	-1	[0.52]	0	[0.51]	6	[0.37]	19	[0.15]	-2	[0.53]	-1	[0.52]	0	[0.50]
FOR	24	[0.12]	33	[0.05]	52	[0.01]	20	[0.12]	21	[0.11]	21	[0.11]	26	[0.07]	33	[0.03]	45	[0.01]	22	[0.07]	22	[0.07]	20	[0.09]
GSI	-17	[0.83]	-12	[0.76]	-4	[0.58]	-1	[0.53]	-1	[0.52]	0	[0.51]	-8	[0.70]	-5	[0.62]	2	[0.46]	2	[0.45]	2	[0.45]	2	[0.44]
INT	8	[0.35]	27	[0.09]	65	[0.00]	0	[0.50]	2	[0.46]	7	[0.37]	18	[0.15]	30	[0.04]	54	[0.00]	2	[0.46]	3	[0.43]	6	[0.37]
JPM	11	[0.27]	28	[0.07]	61	[0.00]	29	[0.09]	33	[0.07]	42	[0.03]	14	[0.22]	25	[0.08]	48	[0.00]	31	[0.07]	34	[0.05]	42	[0.02]
KEM	-1	[0.52]	25	[0.08]	77	[0.00]	-1	[0.52]	17	[0.14]	53	[0.00]	8	[0.30]	25	[0.05]	58	[0.00]	8	[0.28]	19	[0.09]	41	[0.00]
MLI	-1	[0.51]	11	[0.26]	35	[0.02]	-6	[0.62]	-2	[0.55]	4	[0.42]	5	[0.39]	12	[0.24]	25	[0.06]	1	[0.48]	3	[0.43]	6	[0.35]
MSI	18	[0.15]	30	[0.04]	54	[0.00]	5	[0.40]	7	[0.35]	12	[0.27]	21	[0.08]	29	[0.03]	47	[0.00]	5	[0.39]	6	[0.36]	8	[0.32]
NEO	-4	[0.58]	8	[0.35]	32	[0.06]	5	[0.40]	13	[0.26]	28	[0.08]	6	[0.38]	13	[0.23]	29	[0.05]	12	[0.27]	17	[0.18]	28	[0.07]
NIP	-2	[0.53]	4	[0.42]	16	[0.22]	20	[0.19]	21	[0.19]	22	[0.18]	12	[0.26]	17	[0.19]	26	[0.09]	28	[0.10]	27	[0.11]	26	[0.12]
NON	14	[0.26]	21	[0.18]	34	[0.07]	29	[0.11]	28	[0.12]	27	[0.14]	27	[0.10]	31	[0.07]	39	[0.03]	30	[0.08]	27	[0.11]	21	[0.18]
NRD	23	[0.11]	35	[0.03]	58	[0.00]	30	[0.06]	29	[0.07]	27	[0.08]	26	[0.08]	35	[0.03]	53	[0.00]	31	[0.04]	31	[0.04]	31	[0.04]
OPS	-20	[0.79]	-5	[0.58]	24	[0.17]	-13	[0.72]	-6	[0.60]	8	[0.35]	1	[0.48]	11	[0.29]	31	[0.06]	-2	[0.54]	2	[0.46]	11	[0.29]
RBN	8	[0.36]	16	[0.22]	33	[0.06]	16	[0.23]	21	[0.17]	30	[0.09]	13	[0.26]	16	[0.21]	23	[0.13]	22	[0.15]	23	[0.14]	25	[0.12]
SAB	8	[0.35]	14	[0.24]	26	[0.09]	13	[0.26]	14	[0.23]	17	[0.19]	11	[0.28]	16	[0.19]	26	[0.08]	15	[0.22]	16	[0.20]	18	[0.18]
SGP	-11	[0.73]	-4	[0.60]	8	[0.32]	-6	[0.63]	-7	[0.65]	-8	[0.69]	1	[0.49]	3	[0.41]	9	[0.28]	5	[0.37]	4	[0.41]	0	[0.49]
SHB	22	[0.16]	29	[0.09]	44	[0.02]	10	[0.34]	14	[0.27]	25	[0.15]	25	[0.10]	30	[0.06]	40	[0.02]	14	[0.25]	18	[0.19]	26	[0.10]
SWB	15	[0.21]	32	[0.04]	68	[0.00]	41	[0.03]	53	[0.01]	76	[0.00]	22	[0.09]	34	[0.02]	59	[0.00]	43	[0.01]	52	[0.00]	70	[0.00]
UB	11	[0.30]	22	[0.15]	43	[0.02]	14	[0.26]	16	[0.23]	20	[0.18]	13	[0.24]	20	[0.15]	32	[0.04]	16	[0.20]	17	[0.19]	18	[0.18]
UBS	3	[0.44]	16	[0.17]	43	[0.01]	16	[0.17]	16	[0.17]	17	[0.16]	14	[0.20]	21	[0.11]	35	[0.02]	22	[0.06]	22	[0.07]	20	[0.09]
ANON	0.00	0.50	0.00	0.50	0.00	0.50	0.00	0.50	0.00	0.50	0.00	0.50	0.00	0.50	0.00	0.50	0.00	0.50	0.00	0.50	0.00	0.50	0.00	0.50

Table 9: Is Predictive Power Robust to Transaction Costs?.

The table presents the impact of transaction costs on the performance of the mean variance portfolios. The estimation period is 08/10/2010–02/28/2011 (124 days). There is one portfolio for each broker, while the portfolio that disregards the broker identity (ANON) is the benchmark portfolio. Same as before, the investment scenario is based on an annual target volatility $\sigma_p = 10\%$, and daily rebalancing, using the order flow models $M1 - M4$, described in section 2. This recursive regression (out-of-sample) estimation is based on a window of expanding size. We present the annualized Θ for 3 levels of transaction costs (TC); Θ_0 for $TC = 0bps$, Θ_{10} for $TC = 10bps$, and Θ_{30} for $TC = 30bps$. Θ is the difference between the broker^j's and ANON's performance measure (MPPM) of Goetzmann et al. (2007), which is expressed in percentage points and is for $\gamma = 6$. Following Goetzmann et al. (2007), we test whether the broker^j's portfolio significantly outperforms the one that disregards the broker identity and report the p-values in square brackets.

	M1						**M2**						**M3**						**M4**					
	Θ_0	*p-val*	Θ_{10}	*p-val*	Θ_{30}	*p-val*	Θ_0	*p-val*	Θ_{10}	*p-val*	Θ_{30}	*p-val*	Θ_0	*p-val*	Θ_{10}	*p-val*	Θ_{30}	*p-val*	Θ_0	*p-val*	Θ_{10}	*p-val*	Θ_{30}	*p-val*
AAL	10	[0.29]	18	[0.18]	32	[0.05]	6	[0.38]	7	[0.37]	8	[0.34]	4	[0.42]	10	[0.30]	22	[0.12]	2	[0.46]	3	[0.43]	6	[0.38]
BPP	2	[0.44]	3	[0.41]	6	[0.36]	7	[0.33]	0	[0.51]	-16	[0.83]	3	[0.43]	4	[0.38]	8	[0.30]	12	[0.22]	5	[0.37]	-8	[0.70]
CAR	34	[0.01]	36	[0.01]	41	[0.00]	14	[0.20]	12	[0.24]	8	[0.33]	30	[0.01]	33	[0.01]	38	[0.00]	18	[0.13]	17	[0.14]	15	[0.17]
CDG	23	[0.14]	25	[0.11]	31	[0.07]	22	[0.11]	14	[0.22]	-1	[0.53]	23	[0.13]	26	[0.10]	32	[0.06]	21	[0.11]	15	[0.19]	4	[0.41]
CDV	23	[0.08]	35	[0.01]	60	[0.00]	17	[0.17]	18	[0.15]	20	[0.14]	23	[0.07]	32	[0.02]	51	[0.00]	17	[0.16]	19	[0.14]	22	[0.11]
CSB	12	[0.23]	14	[0.19]	18	[0.13]	23	[0.09]	15	[0.20]	-3	[0.56]	6	[0.34]	11	[0.25]	19	[0.12]	22	[0.10]	16	[0.18]	3	[0.42]
DBL	-4	[0.59]	3	[0.43]	17	[0.16]	-8	[0.67]	-7	[0.65]	-5	[0.60]	-5	[0.62]	1	[0.47]	13	[0.21]	-4	[0.59]	-3	[0.56]	0	[0.50]
DDB	8	[0.31]	8	[0.32]	7	[0.33]	14	[0.20]	7	[0.34]	-7	[0.64]	14	[0.19]	15	[0.17]	18	[0.13]	19	[0.13]	14	[0.20]	5	[0.38]
DIF	2	[0.45]	-3	[0.56]	-12	[0.77]	6	[0.36]	-2	[0.54]	-19	[0.85]	6	[0.35]	2	[0.44]	-5	[0.63]	11	[0.27]	4	[0.41]	-10	[0.70]
ENS	15	[0.12]	21	[0.05]	32	[0.01]	-1	[0.53]	-8	[0.68]	-21	[0.89]	18	[0.07]	23	[0.03]	32	[0.00]	3	[0.43]	-3	[0.56]	-14	[0.78]
EVL	20	[0.15]	28	[0.08]	43	[0.01]	16	[0.19]	14	[0.22]	10	[0.29]	19	[0.16]	25	[0.09]	39	[0.02]	18	[0.15]	18	[0.15]	17	[0.17]
FIM	8	[0.34]	9	[0.32]	11	[0.28]	11	[0.27]	8	[0.35]	0	[0.51]	9	[0.31]	10	[0.29]	12	[0.26]	13	[0.24]	9	[0.30]	2	[0.45]
FOR	27	[0.04]	29	[0.03]	33	[0.02]	14	[0.17]	10	[0.25]	2	[0.46]	26	[0.05]	28	[0.03]	32	[0.02]	15	[0.14]	12	[0.19]	6	[0.33]
GSI	20	[0.09]	26	[0.04]	37	[0.01]	17	[0.15]	13	[0.21]	6	[0.36]	22	[0.07]	27	[0.03]	36	[0.01]	22	[0.08]	20	[0.11]	14	[0.19]
INT	18	[0.13]	29	[0.03]	52	[0.00]	17	[0.14]	18	[0.12]	21	[0.10]	18	[0.12]	26	[0.04]	42	[0.00]	18	[0.12]	20	[0.11]	23	[0.08]
JPM	33	[0.04]	42	[0.01]	59	[0.00]	13	[0.24]	9	[0.31]	1	[0.48]	33	[0.04]	40	[0.01]	54	[0.00]	15	[0.20]	12	[0.25]	7	[0.36]
KEM	19	[0.08]	44	[0.00]	92	[0.00]	14	[0.18]	20	[0.10]	30	[0.03]	18	[0.09]	33	[0.01]	64	[0.00]	11	[0.23]	13	[0.18]	19	[0.11]
MLI	8	[0.32]	15	[0.18]	29	[0.04]	7	[0.33]	6	[0.35]	5	[0.39]	7	[0.34]	12	[0.22]	23	[0.08]	11	[0.25]	11	[0.25]	11	[0.26]
MSI	14	[0.21]	25	[0.08]	46	[0.01]	13	[0.27]	14	[0.25]	16	[0.23]	17	[0.17]	25	[0.08]	41	[0.01]	16	[0.21]	17	[0.19]	20	[0.15]
NEO	13	[0.17]	17	[0.11]	24	[0.04]	14	[0.20]	13	[0.21]	12	[0.24]	16	[0.12]	19	[0.08]	25	[0.03]	15	[0.17]	15	[0.18]	14	[0.20]
NIP	3	[0.42]	3	[0.42]	3	[0.42]	0	[0.49]	-2	[0.54]	-6	[0.63]	2	[0.45]	3	[0.44]	4	[0.42]	6	[0.36]	3	[0.42]	-2	[0.54]
NON	-10	[0.73]	-15	[0.81]	-25	[0.92]	-15	[0.80]	-25	[0.92]	-44	[0.99]	-7	[0.68]	-11	[0.76]	-18	[0.87]	-13	[0.78]	-21	[0.89]	-36	[0.98]
NRD	29	[0.02]	34	[0.01]	43	[0.00]	31	[0.03]	26	[0.06]	15	[0.20]	29	[0.02]	34	[0.01]	45	[0.00]	27	[0.06]	23	[0.09]	14	[0.21]
OPS	9	[0.29]	15	[0.18]	27	[0.06]	14	[0.22]	11	[0.27]	5	[0.39]	13	[0.20]	19	[0.11]	31	[0.03]	18	[0.15]	17	[0.17]	15	[0.21]
RBN	36	[0.03]	44	[0.01]	61	[0.00]	38	[0.03]	40	[0.03]	44	[0.02]	38	[0.02]	45	[0.01]	58	[0.00]	43	[0.02]	45	[0.01]	49	[0.01]
SAB	15	[0.22]	21	[0.14]	34	[0.04]	16	[0.19]	14	[0.22]	10	[0.29]	13	[0.24]	20	[0.15]	33	[0.05]	14	[0.22]	13	[0.23]	12	[0.25]
SGP	8	[0.31]	10	[0.27]	14	[0.20]	28	[0.07]	22	[0.13]	10	[0.31]	8	[0.30]	8	[0.31]	7	[0.34]	27	[0.07]	22	[0.12]	11	[0.28]
SHB	29	[0.03]	32	[0.02]	38	[0.01]	23	[0.09]	21	[0.12]	15	[0.20]	30	[0.02]	31	[0.02]	33	[0.01]	26	[0.05]	24	[0.07]	20	[0.11]
SWB	36	[0.03]	46	[0.01]	68	[0.00]	29	[0.06]	29	[0.05]	31	[0.05]	36	[0.02]	45	[0.01]	62	[0.00]	28	[0.07]	29	[0.06]	30	[0.06]
UB	8	[0.29]	12	[0.21]	20	[0.10]	12	[0.23]	8	[0.31]	1	[0.49]	14	[0.17]	16	[0.13]	21	[0.08]	18	[0.12]	14	[0.18]	8	[0.32]
UBS	22	[0.10]	32	[0.03]	52	[0.00]	29	[0.08]	28	[0.08]	27	[0.10]	20	[0.12]	27	[0.06]	42	[0.01]	26	[0.10]	27	[0.09]	28	[0.08]

Table 10: Brokers' Market Share Statistics.

The table presents the market share statistics of brokers in Helsinki Stock Exchange in the period 03/29/2010-02/28/2011. The first statistic (*Aggr. Vol*) is the average daily volume initiated by each broker across the 15 most liquid stocks of HEX25 index. The second statistic (*Vol*) is the average daily volume (aggressive and passive) executed by each broker. Finally, the last column presents the fraction of average daily aggressive trading to average daily total volume.

Broker	*Aggr. Vol*	*Vol*	*Fraction*
AAL	10,808	28,588	0.38
BPP	90,800	186,303	0.49
CAR	39,095	97,922	0.40
CDG	230,219	258,559	0.89
CDV	37,343	80,520	0.46
CSB	172,235	311,911	0.55
DBL	78,616	211,308	0.37
DDB	94,088	163,288	0.58
DIF	10,273	15,539	0.66
ENS	181,505	401,946	0.45
EVL	24,690	55,228	0.45
FIM	91,111	185,570	0.49
FOR	139,010	317,786	0.44
GSI	54,112	118,673	0.46
INT	16,630	36,652	0.45
JPM	35,708	78,671	0.45
KEM	8,940	13,624	0.66
MLI	61,640	135,947	0.45
MSI	69,306	208,440	0.33
NEO	22,368	43,898	0.51
NIP	96,064	134,446	0.71
NON	88,597	194,032	0.46
NRD	112,786	248,048	0.45
OPS	51,368	134,973	0.38
RBN	19,561	44,734	0.44
SAB	62,674	137,816	0.45
SGP	139,975	204,438	0.68
SHB	53,985	149,754	0.36
SWB	37,522	87,252	0.43
UB	13,805	48,032	0.29
UBS	44,406	121,124	0.37
ANON	2,386,043	4,919,366	0.50

Table 11: Do Large Brokers (Q4) Outperform Small (Q1)?

The table presents the performance difference between a portfolio that tracks large brokers and one that tracks small brokers. We use 3 size criteria to categorize brokers: a. the average daily volume initiated (*Aggr. Vol*), b. the average daily volume executed (*Vol*) and c. the ratio of the first 2 criteria (*Fraction*). Next, we construct daily average order flow measure series of the top (Q4) and bottom (Q1) quartile of brokers for each criterion. We build daily rebalancing mean-variance portfolios using the order flow models $M1 - M4$, described in section 2, to predict next day's returns. This out-of-sample recursive regression estimation is based on a window of expanding size in the period 08/10/2010–02/28/2011 (124 days). We estimate the MPPM of Goetzmann et al. (2007) for each group of brokers and report $\Delta\Theta$, which is the performance difference between the two portfolios expressed in annual percentage points and for $\gamma = 6$. Following Goetzmann et al. (2007), we test whether the large brokers' portfolio significantly outperforms ($\Delta\Theta > 0$) the small brokers' portfolio and report the p-values in square brackets. *std* is the standard deviation of the $\Delta\Theta$.

	M1	M2	M3	M4
a. Aggr. Vol				
$\boldsymbol{\Delta\Theta}$	-34	-42	-30	-43
std	18	17	18	16
p-val	[0.03]	[0.01]	[0.05]	[0.00]
b. Vol				
$\boldsymbol{\Delta\Theta}$	-51	-68	-52	-73
std	21	18	21	18
p-val	[0.01]	[0.00]	[0.01]	[0.00]
c. Fraction				
$\boldsymbol{\Delta\Theta}$	-3	-11	-7	-11
std	18	17	19	17
p-val	[0.45]	[0.26]	[0.36]	[0.25]

Table 12: Investor Heterogeneity, Market Share, and Portfolio Performance.

In panel a., the table presents the correlated trading statistics of the daily trades of brokers in HEX in the period 03/29/2010-28/02/2011. We measure correlated trading by the herding measure (LSV) of Lakonishok et al. (1992), which is defined in Equation 11. The LSV statistics are computed for each stock-day and then averaged. If trades are independent, the mean LSV measure will be zero. In panel b., we present the average size statistics of the quartile of brokers with the highest (*High*) and lowest (*Low*) LSV statistic. The size is measured with respect to the average daily volume initiated (*Aggr. Vol*) or executed (*Vol*, aggressive and passive) by each broker. In panel c., we construct daily average order flow measure series of the top (Q4) and bottom (Q1) quartile of brokers, and then we build daily-rebalancing mean-variance portfolios, using the order flow models $M1 - M4$, described in section 2, to predict next day's returns. This out-of-sample recursive regression estimation is based on a window of expanding size in the period 08/10/2010–02/28/2011 (124 days). We estimate the performance measure of Goetzmann et al. (2007) of each quartile of brokers and report $\Delta\hat{\Theta}$, which is the performance difference between the brokers with high (Q4) and low (Q1) correlated trading, expressed in annual percentage points and for $\gamma = 6$. Following Goetzmann et al. (2007), we test whether the top quartile of brokers significantly outperforms ($\Delta\Theta > 0$) the bottom quartile. We report the standard deviation (*std*) of $\Delta\Theta$, and the p-values in square brackets.

a. LSV Statistics

Brokers	*LSV*	*Brokers*	*LSV*	*Brokers*	*LSV*
AAL	0.10	***FIM***	0.05	***NRD***	0.11
BPP	0.17	***FOR***	0.05	***OPS***	0.10
CAR	0.18	***GSI***	0.15	***RBN***	0.27
CDG	0.02	***INT***	0.23	***SAB***	0.14
CDV	0.23	***JPM***	0.24	***SGP***	0.08
CSB	0.09	***KEM***	0.23	***SHB***	0.16
DBL	0.13	***MLI***	0.11	***SWB***	0.14
DDB	0.09	***MSI***	0.19	***UB***	0.14
DIF	0.03	***NEO***	0.12	***UBS***	0.18
ENS	0.15	***NIP***	0.15		
EVL	0.20	***NON***	0.05		

b. LSV and Brokers' Size

LSV	*Aggr Vol*	*Vol*
High (Q4)	32,073	79,874
Low (Q1)	120,689	206,390

c. LSV and Mean-Variance Performance

	M1	M2	M3	M4
$\boldsymbol{\Delta\Theta}$	26	25	28	27
std	19	22	19	22
p-val	[0.08]	[0.13]	[0.07]	[0.11]

Table 13: Does Volume *per se* Convey Information?

The table presents the performance of a mean-variance portfolio that uses the average order flow measure of the most active brokers (Q4 quartile) at time t to predict returns at time $t+1$ (Panel a.), $t+2$ (Panel b.), $t+3$ (Panel c.), and $t+4$ (Panel d.) using the order flow models $M1 - M4$, described in section 2. We rebalance portfolio's weights on a daily frequency. This out-of-sample recursive regression estimation is based on a window of expanding size in the period 08/10/2010–02/28/2011 (124 days). We estimate the MPPM of Goetzmann et al. (2007) and report the performance difference, Θ, against the portfolio that disregards the broker identity (ANON). Θ is expressed in (annualized) percentage points and is for $\gamma = 6$. When $\Theta > 0$, market transparency yields positive economic value to mean-variance investors. Following Goetzmann et al. (2007), we test whether the portfolio significantly outperforms the ANON portfolio and report the p-values in square brackets.

	M1	M2	M3	M4
a. Use OF_t^i ***to predict*** R_{t+1}^i				
Θ	28	39	26	40
p-val	[0.00]	[0.00]	[0.00]	[0.00]
b. Use OF_t^i ***to predict*** R_{t+2}^i				
Θ	7	3	11	6
p-val	[0.30]	[0.39]	[0.18]	[0.31]
c. Use OF_t^i ***to predict*** R_{t+3}^i				
Θ	10	8	10	11
p-val	[0.29]	[0.30]	[0.28]	[0.23]
d. Use OF_t^i ***to predict*** R_{t+4}^i				
Θ	20	15	17	15
p-val	[0.09]	[0.20]	[0.12]	[0.18]

Table 14: Analysis of the Investment Style of Brokers.

The table presents the fraction of positive buy ratio differences across brokers (including the ANON portfolio) for the period 03/29/2010–02/28/2011. We follow Grinblatt and Keloharju (2000) to construct the buy ratios: buy volume/(buy volume+sell volume). In our calculations we use only aggressive trades. Each buy ratio difference is generated by subtracting the average buy ratio of stocks in the bottom quartile (losers) from the average buy ratio of stocks in the top quartile (winners). We use hourly and daily buy ratios, while the past returns used for ranking the stocks are based on the previous hour and day, respectively. We present the fraction of positive buy ratio differences ($BRDif$). Under the hypothesis of no momentum or contrarian behavior, the average buy ratio difference should be zero, and the aforementioned fraction equal to 0.50. A fraction which is larger than 0.50 indicates a momentum trading behavior, while a fraction smaller than 0.50 indicates a contrarian behavior. In square brackets we report the p-values of the standard binomial test ($p - val$) of whether the fraction of buy ratio differences is 0.50, together with the AR(1) adjusted p-values ($p - valadj$) suggested by Grinblatt and Keloharju (2000). We drop zero buy ratio differences from the fraction calculation.

	1 hour			1 day		
Broker	*BRDif*	*p-val*	*p-val adj*	*BRDif*	*p-val*	*p-val adj*
AAL	0.49	[0.45]	[0.45]	0.39	[0.00]	[0.00]
BPP	0.49	[0.59]	[0.59]	0.50	[0.90]	[0.89]
CAR	0.49	[0.59]	[0.59]	0.50	[0.95]	[0.95]
CDG	0.60	[0.00]	[0.00]	0.55	[0.13]	[0.15]
CDV	0.57	[0.00]	[0.00]	0.50	[0.89]	[0.89]
CSB	0.43	[0.00]	[0.00]	0.57	[0.04]	[0.04]
DBL	0.42	[0.00]	[0.00]	0.28	[0.00]	[0.00]
DDB	0.42	[0.00]	[0.00]	0.52	[0.47]	[0.47]
DIF	0.46	[0.00]	[0.00]	0.57	[0.03]	[0.03]
ENS	0.43	[0.00]	[0.00]	0.41	[0.01]	[0.01]
EVL	0.49	[0.30]	[0.30]	0.45	[0.15]	[0.17]
FIM	0.40	[0.00]	[0.00]	0.49	[0.74]	[0.74]
FOR	0.58	[0.00]	[0.00]	0.65	[0.00]	[0.00]
GSI	0.55	[0.00]	[0.00]	0.59	[0.01]	[0.01]
INT	0.56	[0.00]	[0.00]	0.58	[0.02]	[0.03]
JPM	0.59	[0.00]	[0.00]	0.55	[0.17]	[0.15]
KEM	0.60	[0.00]	[0.00]	0.56	[0.13]	[0.12]
MLI	0.52	[0.13]	[0.13]	0.49	[0.74]	[0.75]
MSI	0.43	[0.00]	[0.00]	0.41	[0.00]	[0.00]
NEO	0.53	[0.02]	[0.02]	0.57	[0.04]	[0.04]
NIP	0.55	[0.00]	[0.00]	0.51	[0.74]	[0.74]
NON	0.34	[0.00]	[0.00]	0.53	[0.39]	[0.36]
NRD	0.40	[0.00]	[0.00]	0.35	[0.00]	[0.00]
OPS	0.41	[0.00]	[0.00]	0.39	[0.00]	[0.00]
RBN	0.55	[0.00]	[0.00]	0.53	[0.33]	[0.33]
SAB	0.60	[0.00]	[0.00]	0.68	[0.00]	[0.00]
SGP	0.48	[0.04]	[0.04]	0.48	[0.47]	[0.46]
SHB	0.51	[0.38]	[0.38]	0.55	[0.17]	[0.14]
SWB	0.53	[0.04]	[0.03]	0.58	[0.02]	[0.01]
UB	0.47	[0.06]	[0.06]	0.53	[0.44]	[0.46]
UBS	0.58	[0.00]	[0.00]	0.59	[0.00]	[0.00]
ANON	0.52	[0.02]	[0.02]	0.64	[0.00]	[0.00]

Table 15: Do Momentum Brokers Outperform Contrarian?

The table presents the performance of a mean-variance portfolio that uses the average order flow of momentum brokers at time t to predict returns at time t+1 using the order flow models M1 - M4, described in section 2. We repeat for contrarian brokers. We follow Grinblatt and Keloharju (2000) to characterize brokers as momentum or contrarian. Momentum are the brokers with daily buy ratio difference fraction significantly (p-value$<$ 5%) greater than 0.50. Contrarian are the brokers with daily buy ratio difference fraction significantly (p-value$<$ 5%) smaller than 0.50. We rebalance portfolio's weights on a daily frequency. This out-of-sample recursive regression estimation is based on a window of expanding size in the period 08/10/2010–02/28/2011 (124 days). We estimate the MPPM of Goetzmann et al. (2007) and report the performance difference of the two portfolios against the one that disregards the broker identity (ANON). We also report $\Delta\Theta$, which is the performance difference between the momentum and contrarian portfolio. Performance differences are expressed in annual percentage points and are for $\gamma = 6$. Following Goetzmann et al. (2007), we test if the two portfolios significantly outperform ($\Theta > 0$) the ANON portfolio, as well as if the momentum portfolio significantly outperforms ($\Delta\Theta > 0$) the contrarian. We report the p-values in square brackets.

	M1	M2	M3	M4
Momentum	15	30	16	34
p-val	[0.11]	[0.01]	[0.10]	[0.00]
Contrarian	6	2	4	3
p-val	[0.37]	[0.46]	[0.40]	[0.44]
$\boldsymbol{\Delta\Theta}$	9	28	12	31
p-val	[0.29]	[0.04]	[0.23]	[0.03]

Table 16: Stock Picking Ability (Sophistication) of Brokers.

The table presents the fraction of positive buy ratio differences across brokers for the period 03/29/2010–02/28/2011. We follow Grinblatt and Keloharju (2000) to construct the buy ratios: buy volume/(buy volume+sell volume). In our calculations we use only aggressive trades. Each buy ratio difference is generated by subtracting the average buy ratio of stocks with future one- or three-month returns in the bottom quartile (losers) from the average buy ratio of stocks with future one- or three-month returns in the top quartile (winners). Firstly, we present the fraction of positive buy ratio differences ($BRDif$). In the absence of stock picking ability, the average buy ratio difference should be zero, and the aforementioned fraction equal to 0.50. A fraction larger (smaller) than 0.50 means that the stocks brokers buy on a daily basis have a positive (negative) one- or three-months performance, thus, brokers have high (low) stock picking ability. In square brackets we report the p-values of the standard binomial test ($p - val$) of whether the fraction of buy ratio differences is 0.50, together with the AR(1) adjusted p-values ($p - valadj$) suggested by Grinblatt and Keloharju (2000). We drop zero buy ratio differences from the fraction calculation.

	1 month			3 months		
Broker	*BRDif*	*p-val*	*p-val adj*	*BRDif*	*p-val*	*p-val adj*
AAL	0.49	[0.73]	[0.75]	0.46	[0.32]	[0.36]
BPP	0.49	[0.73]	[0.74]	0.53	[0.49]	[0.49]
CAR	0.50	[1.00]	[1.00]	0.51	[0.82]	[0.83]
CDG	0.55	[0.12]	[0.12]	0.49	[0.82]	[0.83]
CDV	0.44	[0.08]	[0.10]	0.49	[0.75]	[0.77]
CSB	0.54	[0.30]	[0.31]	0.49	[0.82]	[0.83]
DBL	0.54	[0.24]	[0.27]	0.46	[0.25]	[0.28]
DDB	0.49	[0.84]	[0.84]	0.49	[0.82]	[0.84]
DIF	0.45	[0.18]	[0.17]	0.46	[0.25]	[0.22]
ENS	0.46	[0.24]	[0.27]	0.49	[0.70]	[0.72]
EVL	0.58	[0.02]	[0.01]	0.59	[0.02]	[0.02]
FIM	0.54	[0.30]	[0.31]	0.50	[0.94]	[0.94]
FOR	0.58	[0.02]	[0.04]	0.54	[0.25]	[0.32]
GSI	0.52	[0.54]	[0.56]	0.47	[0.49]	[0.50]
INT	0.55	[0.20]	[0.22]	0.50	[1.00]	[1.00]
JPM	0.54	[0.21]	[0.23]	0.51	[0.88]	[0.89]
KEM	0.47	[0.48]	[0.49]	0.41	[0.05]	[0.08]
MLI	0.55	[0.15]	[0.14]	0.51	[0.82]	[0.83]
MSI	0.52	[0.54]	[0.56]	0.47	[0.49]	[0.52]
NEO	0.56	[0.10]	[0.10]	0.56	[0.11]	[0.12]
NIP	0.46	[0.19]	[0.23]	0.47	[0.40]	[0.40]
NON	0.51	[0.73]	[0.74]	0.54	[0.32]	[0.32]
NRD	0.46	[0.24]	[0.30]	0.45	[0.20]	[0.22]
OPS	0.45	[0.12]	[0.11]	0.46	[0.25]	[0.27]
RBN	0.49	[0.88]	[0.89]	0.53	[0.42]	[0.43]
SAB	0.60	[0.00]	[0.01]	0.53	[0.49]	[0.53]
SGP	0.48	[0.54]	[0.53]	0.47	[0.49]	[0.48]
SHB	0.57	[0.03]	[0.05]	0.49	[0.70]	[0.72]
SWB	0.51	[0.72]	[0.72]	0.44	[0.15]	[0.12]
UB	0.45	[0.19]	[0.21]	0.40	[0.02]	[0.02]
UBS	0.56	[0.06]	[0.09]	0.53	[0.40]	[0.47]

Table 17: Is Stock Picking Ability (Sophistication) Related to Investment Style?

The table shows the relation between stock picking ability and investment style in the period 03/29/2010–02/28/2011. We follow Grinblatt and Keloharju (2000) to construct buy ratio difference fractions based on future one- and three-months returns in order to measure brokers' stock picking ability. In the absence of stock picking ability, the average buy ratio difference should be zero, and the aforementioned fraction equal to 0.50. A fraction larger (smaller) than 0.50 means that the stocks brokers buy on a daily basis have a positive (negative) one- or three-months performance, thus, brokers have high (low) stock picking ability. We split brokers into two groups; those with high stock picking ability (Q4 quartile), and those with low stock picking ability (Q1 quartile). We, then, follow Grinblatt and Keloharju (2000) to measure brokers' investment style and report the relevant average buy ratio difference fraction ($BRDif$) of each group based on one-day past returns. A fraction which is larger than 0.50 indicates a momentum trading behavior, while a fraction smaller than 0.50 indicates a contrarian behavior. We, also, report the difference of investment styles of the two groups and the associated p-value in square brackets.

Stock Picking Ability	*BRDif*	*Investment Style*
a. 1 month		
High (Q4)	0.57	*Momentum*
Low (Q1)	0.48	*Contrarian*
Q4-Q1	0.09	
p-val	[0.00]	
b. 3 months		
High (Q4)	0.56	*Momentum*
Low (Q1)	0.46	*Contrarian*
Q4-Q1	0.10	
p-val	[0.00]	