The Aggregate Behaviour of Individual Investors

Andrew Jackson[Ψ]

First Draft: 10th October 2001

This Version: 29th July 2003

ABSTRACT

Behavioural models generally require that the investment decisions of irrational investors aggregate in a systematic way. Using a unique Australian dataset of individual investor trades I investigate the plausibility of this assumption. I find that aggregate individual investor trades do indeed exhibit strong systematic patterns, including negative feedback trading and substantial persistence. In addition the weekly cross-sectional net trades of a large number of independent retail brokerage firms are contemporaneously correlated to a remarkable extent. Thus the aggregation assumption appears plausible.

However I do *not* find that the net trades of retail investors consistently predict future returns in a negative fashion. In fact over the period 1991-2002, the net trades of full-service brokerage clients actually *positively* forecast future short-term market and cross-sectional returns. While small investors do act in a highly systematic fashion, their actions may not, at least in the short run, be classed as irrational.

JEL classification: G10, G12.

Keywords: Individual investors; Internet vs Full-service; Feedback Trading; Herding.

Ψ London Business School, Regents Park, London NW14SA, UK. Telephone + 44 (0) 20 7262 5050. Email arjackson@london.edu. I would like to thank Jan Mahrt-Smith, Francisco Gomes, Viral Acharya, Jeff Harris, Stefan Nagel, David Goldreich, Elroy Dimson, Henri Servaes, seminar participants at the Western Finance Association 2002, Transatlantic PhD Conference 2002, London Business School 2001, and especially Tim Johnson for valuable comments and suggestions. I am responsible for any errors.

I. Introduction

The ability of behavioural models in finance to explain asset returns rests on three key assumptions. Firstly, a group of investors must deviate from rationality in their trading activity. Secondly these errors must be sufficiently systematic across the group to avoid them being diversified away upon aggregation. Thirdly there must be some limits to arbitrage that prevent rational investors immediately correcting any resulting mispricing. While much literature has addressed various deviations from rationality and limits to arbitrage[1], to date relatively little empirical research has focussed on the second assumption. That is, whether individual investor's deviations from rationality (ie errors) are sufficiently systematic to aggregate in a meaningful way.

The aggregation assumption plays a key role in models where the existence of a group of commonly acting irrational investors (or equivalently a representative irrational agent) is evoked[2]. Shleifer (2000) argues that errors *should* aggregate based on the systematic nature of the behavioural biases demonstrated by Kahneman & Tversky (1982) and others. However, since much of this work is experimentally based, it is important to examine this assertion in real-life financial settings where agents typically face strong economic incentives.

This paper examines the aggregation assumption using a unique database of individual investor trades. The database is somewhat larger than that used in previous studies of small investors, containing data from 47 full-service and 9 internet brokerage firms in Australia over an 11-year period from 1991 to 2002[3]. Additional information from 16 institutional and 15 mixed brokerage firms (dealing with both retail & institutional clients) is also used. Firstly I examine the trading behaviour of a large group of individual investors to assess whether there are any systematic patterns in their trading that remain after aggregation. Secondly I examine whether the actions of unrelated subgroups of individual investors from a large number of independent brokerage firms are positively correlated. Finally I examine the relationship between aggregated trades and future returns. The analysis is performed at two levels, at the market level (examining flows into and out of the equity market as a whole) and at the cross-sectional level (examining flows into and out of

---

[1] For a survey of deviations from rationality see Barberis & Thaler (2001), for limits to arbitrage see Shleifer (2000).

[2] See for example DeLong et al (1990) where correlated noise traders cause prices to deviate from their true values, Hong & Stein (1999) where a group of boundedly rational momentum traders exist who trade based on past returns and Baker & Stein (2002) where a group of irrational investors exist who underreact to the information contained in order flow.

[3] This paper uses information from over 41 million small investor trades. By comparison Barber & Odean (2000) use around 3 million trades & Grinblatt & Keloharju (2001) use 2 million trades. Note however that both these papers have account level information, while this paper uses information aggregated at the individual broker level.

individual stocks). Additionally the actions of individuals using both internet and traditional full-service brokers are compared.

Why individuals? Of all investor groups, individual investors seem a priori a likely candidate for a group that might be subject to systematic behaviourally induced errors. For example previous research shows that individuals appear to trade excessively, exercise options in a clearly irrational manner, sell stocks that outperform those they purchase, and hold undiversified stock portfolios with excessive allocations to their employers stock[4]. Additionally as a group, individuals are large, holding 42% of all outstanding equity in the US, 16% in the UK, 20% in Germany and 24% in Australia. Individuals are also likely to have access to fewer resources relative to institutional investors when making their investment decisions. Some individual investors reduce this problem via trading through full-service brokerage firms where they receive investment advice. Others trade via internet or discount brokers where they receive none.

The first section of the paper finds a number of strong systematic patterns in behaviour that survive after aggregation across investors from all 56 retail brokerage firms. To reduce the possibility of unintentional data-mining I limit my analysis to four groups of explanatory variables identified by previous small investor research; namely past returns & volatility, calendar effects and lagged investor flows. In addition I conduct numerous robustness checks across sub-periods and across subsets of stocks. The strongest pattern I find is that individual investors appear to engage in significant levels of negative feedback trading (buying after falls, selling after rises) at *both* the market and cross-sectional levels. This result is consistent with previous studies of small investors at the cross-sectional level (Odean (1998), Grinblatt et al (2000)) however at the aggregate market level it differs from the results in the mutual fund flow literature (Edelen & Warner (2001)). Negative feedback trading is most pronounced for internet brokerage clients, but is also significant for full-service retail brokerage clients. I also find that the net flows of individual investors are extremely persistent over time especially at the cross-sectional level. Furthermore, individuals are net buyers after high individual stock volatility, and individual trading activity is high in the first half of January and in the second half of June (before tax-year end).

In the second section, I examine small investor trades disaggregated to the individual broker level to assess the similarity of actions of small investors from independent brokerage firms. The results show that individual investor actions are highly correlated between different brokerage firms. For example the average pairwise cross-sectional correlation for weekly net flows into the top

[4] See Barber & Odean (2000), Poteshman & Serbin (2003), Odean (1998) and Benartzi (2001), respectively.

50 stocks between unrelated pairs of internet brokerage firms is 0.439. For full-service firms the average pairwise correlation is 0.235. In contrast for unrelated pairs of institutional brokers the correlation is zero. The correlation result for retail brokers is remarkably robust with significant positive correlation coefficients found for *every* single unique pair of retail brokerage firms over the sample period. This result is striking given there are 1081 unique retail broker pairs where overlapping data is available. Concurrent related research by Barber, Odean & Zhu (2003) shows that investors *within* a *single* broker have strongly correlated trades. While correlation within a single broker may be due to common investment advice, or common social networks, it is highly unlikely that such explanations would be valid for the current paper's result that positive correlations exist *across* 56 unrelated brokerage firms with millions of retail clients. As such the results here provide strong evidence that the aggregation assumption is likely to hold.

The final section of the paper examines the relationship between individual investor flows and future stock returns at the market and cross-sectional levels. Vector autoregression analysis is used to allow for joint endogeneity between returns and flows. At the aggregate market level I find small investor flows positively predict future 1-week market returns with evidence of subsequent reversals, consistent with short-term price pressure. At the cross-sectional level, I find small investors predict futures 2-3 week excess returns with no evidence of reversals, consistent with private information or liquidity provision. Further examination reveals that this result is entirely driven by full-service brokerage clients. In a related finding, Pan & Poteshman (2003) show that full-service brokerage clients' equity option trades positively forecast future stock returns. The results do not support the conjecture that behavioural errors are causing individual investors to act in an irrational manner at either the aggregate market or cross-sectional level.

In summary, the main contribution of this paper is to provide strong evidence that the aggregation assumption *does* appear to hold for small investors. Aggregated flows from this group of investors *do* contain strong systematic patterns and the actions of individual investors from a large number of unrelated retail brokerage firms are highly correlated. The final section of the paper reveals however that behavioural models that assume small investors are in aggregate irrational or uninformed may not be justified, given that the net flows of full-service investors *positively* predict future short-term cross-sectional and aggregate market returns.

The paper is structured as follows. Section II describes the data used in the study. Section III examines aggregated small investor flows for evidence of systematic patterns. Section IV examines correlations in net flows across a large number of brokerage firms. Section V examines the relationship between net flows and future returns. Section VI concludes.

II. Data Description

This paper uses a unique dataset of small investor trading on the Australian Stock Exchange. The dataset contains information on the trades of fifty-six retail brokers (nine internet brokers and forty-seven traditional full-service brokers) over the period September 1991 to December 2002[5][6]. For each of the fifty-six brokers, data is available on the number of trades, the value traded, and the average price of all buy and sell trades executed during each week, for each stock. The dataset is proprietary and was supplied by a leading Australian broker. Trading on the Australian Stock Exchange during this period was conducted using the automated SEATS system that operates via a fully transparent electronic open limit order book with no designated market makers[7]. In this system all trades are reported and tagged with a unique broker identification code that enables us to track the specific broker engaged on each side of each transaction. The broker identification data used in this study was observable in real-time on a trade-by-trade basis by some market participants (brokers) but was not available to investors in general.

Summary statistics describing the dataset appear in Table 1. In aggregate around 21.6 million buy and 20.2 million sell transactions are represented with an aggregate value of $A191 billion and $A188 billion respectively. Additionally the sample is split into two groups[8], internet brokerage investors and traditional full-service brokerage investors. Table 1 shows that, on average, the clients of these retail brokers trade small amounts, with an average transaction size of $7,570 for internet brokerage investors and $10,080 for full-service brokerage investors[9]. As found in Barber & Odean (2000), the average value of stocks sold is slightly higher than the average value of stocks purchased ($9,292 vs $8,824) and the number of buy transactions is higher than the number of sell transactions.

---

[5] Not all brokers exist for the full period. For example most internet brokers start after 1995, while several brokers leave the sample over the 11-year period.

[6] Additional information on 16 institutional brokers and 15 mixed brokers (that deal with both retail and institutional clients) is also used.

[7] For more details on the institutional structure of the ASX market see Aitken et al (1995). Note that although the trading system does not include a designated market maker, there exists a well established pseudo-market making system as evidenced in Fong et al (2001) who show that for the highest decile volume stocks, 34% of trading is done in the upstairs market (via crossings).

[8] Internet brokers were identified using the ASX 2001 Participants directory and confirmed via examining broker web sites.

[9] Occasionally institutional trades appear in the records for these brokers. To remove the influence of these trades, stock weeks with an average trade size of >$A50000 for either buy or sell transactions were removed and the data for the week was

The fifty-six brokers in the sample account for a significant proportion of all trades executed by small investors over this period. It is difficult to quantify this figure exactly for the full-service portion of the sample, as some small investors trade through full-service mixed brokerage houses that deal with both institutional and small investors. Based on the number of trades from these mixed firms I estimate that the percentage of all full-service retail brokerage trades captured by my 47 full-service brokers is over 50%. In contrast the 9 internet brokers in the sample capture virtually all internet based trades.

The economic significance of the sample can be demonstrated by calculating the total market share (by value traded) of these brokers relative to the total value traded in the entire market. Market share results are shown in Figure 1. The market share of the 47 full-service brokers remains at around 5% throughout the entire period. The market share of the 9 internet brokers grows rapidly from early 1998 to reach around 5% by the end of the sample period[10]. Figure 2 shows that the market share of these brokers calculated by the *number* of trades is significantly higher reaching around one-third of all trades executed in the final three years of the sample.

While the combined aggregate market share of less than 10% may lead us to conclude that this group of investors is likely to have little impact on market dynamics, the market share of these investors in the *typical* index stock is significantly higher. As found in Lee, Shleifer & Thaler (1991), market share of small investors increases as firm size decreases. Table 1B shows the average market share (by value) for stocks in five different size groupings. For large stocks market share of these investors is small, however for the average index stock in the sample market share of these investors is around 25% for the full sample, and around 33% in the most recent three years. This indicates that small investors are of sufficient size to potentially influence market dynamics for the average index stock.

This paper restricts the universe of stocks to those included in the All Ordinaries Index up to April 2000, and stocks in the ASX300 index for the remainder of the period[11]. These stocks represent over 90% of total outstanding market capitalisation over the period, are held by both

treated as missing. The number of records removed was small, 0.15% of Internet broker stock-weeks were removed, and 1.07% of full-service broker trades were removed.

[10] A large spike in market share can be seen in late December 1999 driven by the Y2K effect, with extremely low institutional volume in the week prior to the new millennium.

[11] On the 3rd April 2000 the construction of the All Ordinaries index was changed significantly, with the number of stocks in the index going from around 266 to 500. The ASX300 post April 2000 is the closest market index to the All Ordinaries prior to the changes and became the equivalent institutional benchmark. Source: ASX Factbook 2001.

institutional and individual investors, and are of reasonable liquidity. The number of stocks in the index varies between 250 and 350 for each week of the sample with the total market cap of the index varying between $A160bn and $A680bn. Stocks outside the All Ordinaries suffer from extreme illiquidity and resulting non-synchronous trading problems. As a group these micro-cap stocks are not particularly economically important relative to the market as a whole and are omitted from this study.

Supplementary information on closing prices, volume and market returns were obtained from Datastream and cross-checked against data provided by a large Australian investment management firm[12]. Information on daily index composition was obtained from a brokerage firm and verified against the ASX Monthly Index Almanacs.

Finally, it is interesting to know how potentially generalizable the results in this paper are to other equity markets. Table 1C presents some summary statistics on the ownership composition and size of the Australian market relative to other international markets. Individual investor participation is high in Australia (41%) relative to other countries. However the actual percentage of the market owned by individuals (20%) is similar to levels in Japan, UK & France but lower than US individual ownership (42%). Foreign ownership of Australia equities (24%) is similar to that of Japan & UK, but significantly higher than foreign ownership in the US (11%).

III. Systematic Patterns in Aggregate Individual Trading Behaviour

This section examines whether there are systematic patterns that exist in the aggregated net flows of the fifty-six retail brokers. Since all completed trades have both a buyer and a seller, then across the entire market net flows must sum to zero for any given stock each week. As such, if we take a randomly chosen group of traders and aggregate their trades, then in a regression of aggregate net trades on any factor we choose we should observe a zero loading. For example since net trades sum to zero, not all investors can be momentum investors. If however we aggregate the net flows from a group of investors that we believe a priori may behave in a similar way (here small investors), then we may get significant coefficients if we regress these flows on factors that capture the commonality in behaviour.

---

[12] Where data did not match I performed a third check against independent data from the ASX monthly almanacs and data from a brokerage firm. This process identified several errors in the Datastream series, a list of which are available from the author on request.

Specifically, the next two sub-sections examine the relationship between net flows of small investors and past net flows, current and past returns, calendar effects, and current and past volatility. This is done firstly at the market level (flows into and out of the equity market as a whole) and secondly at the cross-sectional level (flows into and out of individual stocks). In addition buy, sell and total flows are examined individually to further understand the nature of the patterns found.

From a behavioural perspective, if no patterns were found that survived after aggregation, then the crucial second assumption of many behavioural models in finance would be violated for this particular group of investors and choice of regressors. Conversely if patterns in aggregate trades are found, this would support the aggregation assumption. However we would need to ensure that these patterns are truly robust and are not simply the product of unintentional data mining. To limit this possibility I limit my analysis to four groups of explanatory variables identified by previous small investor research; namely past returns & volatility, calendar effects, and lagged investor flows.

A. Market Level Flows

Market level flows represent the total dollar amount injected or removed from the equity market as a whole by small investors each week. Market level buy (sell) flows are calculated by summing the total value of buy (sell) orders across all stocks each week for all retail brokers.

Market level *net flows* are defined as:

$$NetFlowMkt_t = \frac{buyflow_t - sellflow_t}{buyflow_t + sellflow_t}$$

This gives a value between +1 (100% buys) and –1 (100% sells) each week. When analysing the buy, sell and total market flows, I detrend these series to ensure stationarity (see Appendix I for a discussion of the methodology).

Net flows together with buy, sell and total flows are regressed on the following factors:

- Contemporaneous and lagged weekly market returns
- Lagged flow information
- Calendar dummies
- Contemporaneous and lagged intra-week market return volatility

For each flow variable, two regressions are run. The first regression includes contemporaneous return and volatility variables together with four weekly lags for returns, flows and volatility. Two longer-term lagged return and flow measures are also included. The second regression uses a reduced number of regressors, with the specification chosen by a recursive procedure designed to minimise the Akaike Information Criterion (AIC). Table 2 presents the results of the analysis.

The first result from Table 2 is that aggregate market level net flows are significantly *negatively* related to contemporaneous and lagged aggregate market returns. In economic terms the relationship is significant. A 10% positive return leads to a –0.273 contemporaneous change in the net flow variable (equivalent to a –2.46 standard deviation shock). The result contrasts with *mutual fund* flows where previous research finds a *positive* relationship between mutual fund inflows and contemporaneous aggregate market returns (Warther (1995)), and a *positive* relationship between fund flows and lagged short-term returns (Edelen & Warner (2001))[13]. The difference between flows into equity mutual funds and direct equities is somewhat puzzling. However the negative feedback trading result is consistent with the *cross-sectional* results of Grinblatt & Keloharju (2001) and Odean (1998) that find individual investor net flows are negatively related to recent returns[14].

The second result involves lagged flows. Warther (1995) finds that aggregate mutual fund flows are highly persistent. Consistent with this result, we can see from Table 2 that market level net flows are also strongly persistent for the retail brokers in this sample.

The third group of regressors examines calendar effects on flows. Ritter (1988) finds that there is a low buy/sell ratio for individuals in late December and a high ratio in early January due to tax year-end driven trading. The results here show no economically significant effects for net flows at the turn of the year. There is weak evidence that net flows are positive in early July, corresponding to the start of the new tax-year in Australia. A more robust result is that the level of trading activity is strongly influenced by calendar effects. Buy, sell and total flows are significantly lower in the second half of December due the Christmas-New Year holiday effect. Total trading activity is also significantly higher than usual in the second half of June, coinciding with the end of the tax year (30th June) and in the first two weeks of January.

The final group of regressors examines short-term market volatility. Table 2 shows that net flows are unaffected by market volatility while buy, sell and total flows are all strongly increasing

---

[13] Cohen (1999) also finds a positive relationship between individual's allocation to equities and past market returns over longer horizons. He attributes this result to relative risk aversion of households decreasing in wealth.

[14] These latter papers explain this result via the disposition effect of Shefrin & Statman (1985). Alternative rational explanations may include portfolio rebalancing or individuals acting as liquidity providers to the market. For the objectives of this paper we are not concerned whether the patterns are rationally or behaviourally induced, we are simply interested whether systematic patterns exist that survive in a large group of small investors.

in contemporaneous volatility. This matches the well-known relationship between volume and volatility as shown by Karpoff (1987).

Next I examine whether the aggregate flows generated by internet broking clients are significantly different from aggregate flows generated by full-service broking clients. As noted by Odean (1998) the actions of discount and full-service brokerage clients may differ. He uses discount clients in his analysis to avoid the need to disentangle the decisions and motivations of individual investors from those of their retail brokers. In Table 3 both sets of clients are separated to see if there are any significant differences in their aggregated flows.

Net flows, buy, sell and total flows are constructed separately for both internet and full-service brokers. The same de-trending and normalisation process is applied as for the full sample results. For brevity only the results for the AIC optimised regressions are presented, rather than the full regression results. Since the internet sample began only part way through the sample period, the period for the two regressions is from December 1996 to December 2002.

The results of this analysis in Table 3 show that there is little economic difference in the trading behaviour between clients of internet brokers and full-service brokers at the aggregate market level. Both exhibit a negative relationship between returns and net flows, and both show significant positive persistence in one-week aggregate flows. The two groups have similar seasonalities in buy, sell and total flows and also show common behaviour relative to intra-week market volatility.

In summary, based on the results for *market* level flows, it appears that systematic patterns *do* exist in the aggregated flows of individual investors. Moreover, these patterns are consistent across investors from both full-service and internet brokerage firms. The next section examines cross-sectional flows to see if a similar result holds at the individual stock level.

B. Cross Sectional Flows

Cross sectional flows are defined as the aggregated flows from individual investors in and out of individual stocks on a weekly basis. Net flows are calculated using the same definition as for market flows, giving a value between –1 and +1 for each stock in each week of the sample. I also separately model buy, sell and total flows, which are logged and divided by market capitalisation to generate standardised measures for each stock. As discussed earlier, only index stocks are analysed.

Weekly aggregate flows are regressed on the following factors:

- contemporaneous and lagged weekly stock returns

- lagged weekly stock flows
- intra-week stock return volatility

Weekly Fama-Macbeth cross-sectional regressions are used to calculate regression coefficients and Newey-West standard errors are used to adjust for serial dependence in the coefficients caused by the use of overlapping lagged data. In addition contemporaneous and one-week lagged returns are split into positive and negative components to examine any asymmetry in the reaction to positive and negative stock returns.

The results of the regressions using all retail brokers appear in Table 4. The first set of regressors show that weekly net flows are strongly *negatively* related to contemporaneous weekly stock returns for both positive and negative returns. Additionally weekly net flows have a significant negative relationship with lagged stock returns out to a lag of two months. The results support the existence of very strong negative feedback (contrarian) trading at the individual stock level by small investors. This confirms the cross-sectional findings of Odean (1998), Grinblatt & Keloharju (2000) and Shapira & Venezia (2000) that find individuals tend to be contrarian investors[15][16].

Examining the results for buy and sell flows splitting out positive and negative returns reveal some interesting results. Notably, positive returns lead to an increase in both buying and selling by individuals, however sell flows dominate leading to a negative flow in response to positive returns. Negative returns by contrast have little impact on selling behaviour, but have a strong effect on buying behaviour, leading to contrarian behaviour. Why is this interesting? Previous research has explained the negative relationship between net flows and returns as a manifestation of the disposition effect of Shefrin & Statman (1985). The disposition effect implies that investors are reluctant to sell their losers and eager to sell their winners implying net flows will be negatively related to recent weekly returns. However, the disposition effect operates entirely through *selling* behaviour by individuals. The results in Table 5 show that *both* buying and selling behaviour is

[15] Dhar & Kumar (2001) find using account level detail that there is some heterogeneity of small investors in their response to past price trends. As a group though, individual investors appear to behave in a contrarian manner.

[16] Barber, Odean & Zhou (2003) find that over longer horizons (1-3 years) individual investors become positive feedback traders while over shorter horizons (1-6 months) they are contrarian. Repeating my regressions using quarterly net flows rather than weekly net flows provides some support for this finding. In unreported analysis I also find evidence of positive feedback trading over a 12-18 month horizon. However, while the result is statistically significant, the economic magnitude of the long term positive flow-return reaction is significantly weaker than the contrarian trading patterns demonstrated in this section.

driving the negative feedback relationship. The relationship between buys and past returns is not predicted by the disposition explanation[17][18].

The second set of regressors in Table 4 show that net flows are extremely persistent at the cross-sectional level with significant coefficients extending out to two months. High persistence may be indicative of individuals using correlated information sources that are acted on with a lag by some investors.

The third set of regressors in Table 4 shows that there is a *positive* relationship between contemporaneous stock volatility and net, buy, sell and total flows. The total flow – volatility relationship is consistent with the results of Gallant et al (1992) and Epps & Epps (1976) who find that volatility and trading volume are positively related. However, more intriguingly, high volatility generates more buys than sells, resulting in a positive *net* flow on average from individual investors. Barber & Odean (2002) also show that small investors are more likely to trade in stocks that have had recent extreme performance, possibly due to attention effects.

Table 5 separates individuals into full-service and internet brokerage firms to see how similar their behaviour is over the period December 1996 to December 2002. Table 5 reveals that both groups have a negative relationship between contemporaneous stock returns and net flows, however the relationship is significantly *stronger* for internet investors. The economic pattern is still the same however; both groups are strongly contrarian, buy flows strongly increase with contemporaneous negative returns, while sell flows are affected little. Conversely sell flows increase strongly with positive returns while buy flows actually increase slightly.

Both groups exhibit very strong persistence in cross-sectional order flow, with significant coefficients extending out to a two-month lag. Additionally both groups increase trading activity when volatility is high, with increased volatility leading to net buying, especially for the internet investors.

---

[17] An alternative behavioural explanation of the negative flow/return relationship may be anchoring (Kahneman & Tversky (1982)). Shiller (1998) suggests that the past stock price may be an important anchor point for individual investors, and suggests that the more ambiguous the value of the commodity, the more important anchoring is likely to be. If investors were using the past price as an anchor, this would immediately explain the negative return-flow relationship, as after a stock price fall investors would perceive the stock as cheap and vice versa.

[18] Again, the primary objective of this paper is to assess whether there exist systematic patterns that survive aggregatation, not to determine whether they are behaviourally or rationally driven. Negative feedback trading may also be generated by portfolio rebalancing to restore diversification after price movements or by small investors rationally following mean reversion strategies. The important point here is that there are very strong systematic patterns in individual investor flows related to past returns.

We can conclude from this analysis that there *are* strong systematic patterns that exist in the net flows of individual investors at the cross-sectional level. These patterns survive aggregation over a large number of individuals, and are robust to the type of brokerage firm that the individuals trade through.

To ensure these patterns are robust under alternative empirical specifications, I re-estimate Table 4 and Table 5 using the *number* of buys and sells in place of the value of buys and sells. The results are extremely similar, revealing strong contrarian patterns, strong persistence in the number of buys and sells, and net buying in response to increased volatility.

Next I divide the stocks into five size quintiles and estimate the regressions for Table 4 separately for each size group. Flows remain highly persistent across the five size groups. The positive relationship between volatility and net flows is strongest for the largest stocks, and becomes progressively weaker for smaller stocks. The negative feedback trading result is robust across each size quintile. Figure 2 plots the coefficients for contemporaneous and one week lagged positive and negative flows for each size quintile. We can see that the negative feedback results are actually stronger for *larger* stocks[19].

As a third robustness check, I re-estimate Table 4 individually for 49 of the 56 individual retail brokerage firms (those with sufficiently long panels). Of these 49 there are 40 full-service brokers and 9 internet brokers. This analysis finds significant negative coefficients on PosReturn(t) and NegReturn(t) for 47 of the 49 brokers. Significant positive persistence in flows is also present for 47 of the 49 brokers. The volatility result appears to be slightly less robust, with a significant coefficient on contemporaneous volatility found for only 9 of the 49 brokers (5 of these 9 were internet brokers).

As a final robustness check, to show the negative feedback trading pattern is consistent over time I break the sample into yearly blocks from 1992 to 2002 and re-estimate Table 4 separately for each year. Figure 3 plots the estimated coefficients for the contemporaneous and one period lagged return variables. This shows that in *every* single year in the sub-sample there is a negative relationship between net flows and returns at the cross-sectional level. In addition, the t-statistics on these coefficients are large ranging from–1.93 (PosReturn(t-1) in 1992) as the least significant to –17.95 as the most significant (NegReturn(t) in 1999). Moreover in 9 of the 11 subperiods, the

[19] In related research Ranguelova (2001) finds in the US that the disposition effect is strongest for larger stocks, while a reverse disposition effect holds for small stocks. My data confirms that negative feedback trading is stronger for larger stocks and feedback effects have the same sign for all size quintiles.

negative contemporaneous return coefficients are higher than the positive return coefficients. These results confirm our earlier finding that in aggregate individual investors are consistently *contrarian* in their investment decisions, and react more strongly to negative than to positive returns.

To summarise, the results in Section III demonstrate that there *are* systematic patterns in individual investor trades that survive after aggregation over a large number of small investors. These patterns are remarkably consistent over an 11-year time period and are somewhat stronger for larger stocks. In the next section we examine the relationship between small investor flows at a more disaggregated level to assess the similarity of small investor behaviour across a wide range of independent brokerage firms.

IV. Correlations Between Flows from Different Retail Brokerage Firms

If individual investors trades are to aggregate, then the actions of unrelated subgroups of small investors should be positively correlated on average. Contemporaneous research by Barber, Odean & Zhu (2003) shows that the trades of individual investors appear to be correlated *within* single brokerage firms. This result holds for both a single full-service brokerage firm, and a single discount brokerage firm. This section extends this finding to examine whether individual investor trades are systematically contemporaneously correlated *across* independent brokerage firms.

One reason we might find contemporaneous correlation *within* a single firm (especially a traditional full-service firm) is due to correlated investment research provided to the clients of the firm. Alternatively individuals who live in a certain region may use the same brokerage firm, leading to correlated liquidity shocks due to regional economic factors. However if we can show that individual investor transactions are correlated *across* a large number of independent brokerage firms, this would significantly strengthen the result that individual investor trades contain a strong systematic component.

To test this result, an expanded sample of brokers is used, including 16 institutional brokers, 15 mixed brokers (institutional and retail), 47 full-service brokers and 9 internet brokers, giving 87 brokers in total. If individual investors have strong common systematic factors in their trading activity, we should see positive pairwise correlations between the 56 full-service and internet brokers. We should also see positive but weaker correlations for these 56 brokers with the 15 mixed brokers (who deal with some small investors) and zero (or negative) correlations with the institutional brokers.

To test for contemporaneous correlations, I first calculate the cross-sectional net flows each week for each broker for each index stock. Next, for each of the 3741 broker pairs (87x86/2) I calculate the contemporaneous correlation between the net flows of the brokers for each weekly cross-section of between 250 & 350 stocks. Then I take the time series average of this correlation across the full sample and estimate the standard deviation of this series (allowing for autocorrelation in this series induced by persistent flows over time using the Newey West approach). This gives me a point estimate for the correlation and a corresponding t-statistic relative to the null of zero correlation for each individual broker pair. Broker pairs where there are less than 12 overlapping weeks are omitted from the analysis (266 broker pairs or 7.1% of all pairs are omitted). Such excluded pairs arise because some brokers (eg internet brokers) enter partway through the sample, while some brokers exit the sample during the period.

Figure 4 presents a correlation heat map between the net flows of 70 of the 87 brokers[20]. The brokers are sorted into four blocks; 11 institutional, 12 mixed, 39 full-service small brokers & 8 internet small brokers. To further highlight patterns in the data, brokers within each block are sorted in descending order by average trade size for that broker. This implies brokers with smaller clients will be on the right hand side of each block. Figure 4 reveals some interesting results. Looking at the top right corner, we see that the correlation between the 8 internet brokers is high, as is the correlation between the 39 full-service brokers. For the full-service brokers, the correlations rise as the average trade size of the brokerage firm falls, indicating that full-service brokers with smaller clients tend to have more highly correlated net flows. Correlations between pairs of institutional brokers are slightly negative, and correlations between institutional and retail brokerage firms are negative.

Average summary correlations for each group are presented in Table 5. The first four rows of the table report within-group correlations. We can see that the average correlation of weekly cross-sectional net flows for internet broker pairs is 0.229, which is significantly positive. Additionally and remarkably for the 28 unique internet broker pairs, *every* single pair has a positive correlation between cross-sectional net flows that is statistically significant. For full-service brokers, the average correlation is 0.159, and again *every* single broker pair has a positive correlation between net flows that is statistically significant. For the 12 mixed brokers, the average correlation is close to zero, and for the 11 institutional brokers the correlation is slightly negative. For the institutional brokers, 87% of broker-pairs have a statistically significantly negative correlation of net flows.

[20] 17 brokers with a large number of missing pairs are omitted for the figure to improve visualisation of the correlation structure. These brokers are not excluded from the tables.

These results show that individual investor net flows are significantly positively correlated across a large number of different retail brokerage firms. This is not the case for institutional broker net flows.

The second half of Table 5 reports the correlations between pairs of firms in different groups. For the 312 full-service/internet broker pairs the average pairwise correlation is 0.156. Again 100% of these pairs have a positive statistically significant correlation. For the internet/mixed and full-service/mixed pairs the correlations are 0.059 and 0.060 respectively. Finally for broker pairs matched with the institutional brokers, the correlations are negative, with a large percentage of negative (but small) significant pairs.

Since net flows for a given stock each week for the entire market must sum to zero by definition, if we find positive correlations between one group of investor's net flows, they must be negatively related to some other group to satisfy this adding up constraint. It appears that the net flows from small investors are significantly positively correlated, while institutional brokerage firms are negatively correlated with small investor's net trades. Interestingly the correlations of net flows for different institutional brokerage firms are negatively correlated with each other, so systematic common patterns in the trades of their clients are significantly less prevalent (or are internalised within the broker via the matching process).

To test whether the overall correlation results are robust, I repeat the analysis firstly using only the largest 50 index stocks in each week of the sample[21]. The correlation heat map for this test appears in Figure 5. We can see that the correlations are significantly *higher* if we only consider the top 50 stocks. This arises because brokers are more likely to trade in these stocks each week, while for small stocks there are quite a few zero net flow weeks that tend to reduce the correlations. The results for the top 50 stocks appear in Table 7. The average correlation between internet brokerage pairs rises to 0.439, between full-service pairs to 0.270 and between internet/full-service pairs to 0.235. In a second robustness check I re-estimate the pairwise correlations for all brokers and all stocks excluding a subset of widely held large stocks that were privatised or demutualised over the sample period. These 18 stocks[22] were initially allocated disproportionately to small investors who

[21] In additional analysis I calculate the correlation of aggregate market flows between the 87 brokers. I find that the average correlation of aggregate net flows between internet brokers is 0.530, between full-service brokers is 0.153. These are statistically significantly different from zero at the 1% level. This implies that in addition to the correlation of cross-sectional flows, aggregate flows are also highly correlated across different firms.

[22] The specific stocks excluded were Commonwealth Bank (CBA), St George (SGB), Adelaide Bank (ADB), Commonwealth Serum Laboratories (CSL), Tabcorp (TAH), Qantas (QAN), Bankwest (BWA), AMP (AMP), NSW Tab (TAB), Telstra (TLS), Australian Stock Exchange (ASX), Queensland Tab (UTB), AXA (AXA), NRMA (IAG), Colonial (CGH), Suncorp (SUN),

consistently sold down their large initial holdings over the post-listing period, which might lead to correlated trading. The results of this test reveal however that removing these stocks has little impact on the results. The average correlation between internet pairs is 0.217, between full-service pairs is 0.147, and between internet/full-service pairs is 0.143.

These results confirm that the actions of individual investors are strongly systematic *across* a wide range of different brokerage firms. The economic source of this correlation is unlikely to be individual brokerage firm advice, since the investors are coming from 56 separate brokerage firms. The results from section 2 suggest that the correlation in behaviour is more likely driven by a common reaction to recent stock returns and volatility.

The results of these first two sections empirically confirm that the aggregation assumption used in many behavioural finance models may be reasonable for small investors as a group. There appear to be strong systematic patterns in their aggregate trading behaviour, and trades are strongly correlated across different subgroups of small investors.

V. Affect of Aggregate Systematic Patterns on Prices

The final section of this paper examines the relationship between individual investor flows and *future* stock returns to assess whether flows can predict future returns. Previous research on this issue are mixed with some researchers finding that individual investors underperform on average (eg Odean (1998), Barber & Odean (2000)[23]) while others find they outperform (eg Choe, Kho & Stulz (2000)). The relationship between flows and future performance is important as it allows us to assess whether individual investors are acting in an irrationally systematic, or conversely a rationally systematic fashion.

To evaluate the relationship, I treat returns and net flows as jointly endogenous and estimate a vector autoregression (VAR) system at both the aggregate market and cross-sectional levels.

SGIO (SGI) and GIO (GIO). These stocks were identified from the RBA report "Privatisation in Australia" (1997) and from industry reports.

[23] Barber & Odean (2000) provide convincing evidence that small investors after-cost investment performance is generally negative relative to buy & hold strategies. The underperformance appears to be increasing in the level of trading intensity. My analysis focuses entirely on pre-cost performance given the limitations of my data sample.

A. Market Level Flows

In this section I examine the relationship between aggregate market returns and the aggregated market level flows of individual investors. In addition I repeat the analysis separately for full-service and internet brokers. Figure 6A plots the market level aggregated net flows for internet and full-service brokers over time. Figure 6B plots the market accumulation index over the sample period.

To analyse the relationship I estimate the following reduced form VAR system with net flows and aggregate market returns as the dependent variables.

$$z_t = \Phi_0 + \Phi_1 z_{t-1} + \Phi_2 z_{t-2} + \Phi_3 z_{t-3} + \Phi_X x_t + \varepsilon_t$$

$$\text{where: } z_t = [f_t, r_t]^T,\ E(\varepsilon_t \varepsilon_t') = \Omega = \begin{bmatrix} a & c \\ c & b \end{bmatrix}$$

A vector of constants Φ_0 is included, along with a vector of exogenous variables x_t (lagged volatility and calendar dummies). Hamilton (1994) shows that the maximum likelihood estimates of the VAR parameters Φ_i can be found through equation-by-equation OLS estimates of each element on all lags in the system[24]. I estimate the optimal lag length using a likelihood ratio test including the small sample adjustment suggested in Sims (1980). The LR test fails to reject at 5% the null that three lags are significant against various alternative choices[25]. The estimated reduced form VAR specification is presented in Table 8, with the coefficients on the x-variables omitted for brevity. Results are also shown separately for internet and full-service brokerage investors using the shorter period December 1996 to December 2002 for comparability.

The results presented in Table 8 can be interpreted as a forecasting equation and are quite striking. Aggregate market returns are positively related to the previous week net flows of small investors, and negatively related to net flows from the week preceding this. The Granger causality test indicates that the flows significantly granger cause returns with a p-value of 0.02. The effect is economically significant, with a two-standard deviation positive shock to small investor net flows

[24] The MLE estimate of the VAR error covariance matrix is also obtained by using the residuals from the OLS estimation.

[25] The AIC is also minimised at a choice of three lags.

(increasing the net flow ratio by 0.22) increasing the expected market return in the following week by 47bp[26]. We can also see from the VAR that weekly net flows are highly persistent.

Splitting the VAR into full-service and internet brokerage clients, we see from Table 8 that the positive relationship between lagged flows and future returns is predominantly driven by full-service brokerage flows. Additionally the persistence of net flows for these brokers is significantly higher than for internet brokers.

To check whether the flow-returns relationship is robust, I re-estimate the VAR equations separately for each year in the sample assess the stability of the coefficients over time. The flow coefficients in the return prediction equation are shown in Figure 7A. We can see that the coefficient for flow(t-1) is positive for every single year in the sample. The coefficient on flow(t-2) is negative on average. Under the null that the coefficient is zero, the probability of observing 11 positive values is highly significant using a sign test. Additionally I estimate year-by-year coefficient estimates for each type of broker separately (internet, full-service, mixed and institutional). These results are presented in Figure 7B. This shows that in contrast to the full-service broker results, the corresponding coefficients for mixed and institutional brokers are not consistently positive over the sample period. Over the full period, mixed and institutional brokers have a zero (or slightly negative) coefficient for flow(t-1) in predicting future aggregate market returns. Finally I repeat the analysis using excess market returns over the risk-free rate. This results in virtually identical coefficients to those in Table 8.

It appears that the flow – return relationship is reasonably robust, at least for full-service brokerage clients. The reduced form VAR can be used to make forecasts of future returns, however to interpret the VAR economically we need to identify the coefficients in the underlying structural VAR.

$$Az_t = B_0 + B_1 z_{t-1} + B_2 z_{t-2} + B_3 z_{t-3} + B_x x_t + e_t$$

$$\text{where } A = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}, E(e_t e_t') = \Omega = \begin{bmatrix} \sigma_f^2 & 0 \\ 0 & \sigma_R^2 \end{bmatrix}$$

Identification allows us to make statements about the economic effect of a pure shock to one variable (flows or returns) on the system as a whole. Estimating the reduced form VAR does not yield sufficient restrictions on parameters to identify the structural parameters uniquely. The

[26] Including different numbers of lags in the VAR does not alter this result significantly. For example 4 lags gives an equivalent result of +50bp while 2 lags gives +52bp.

standard approach in VAR analysis is to impose the condition that $b_{12} = 0$ or $b_{21} = 0$ (the Choleski decomposition) however if this imposed condition is not true then Enders (1995 p322) notes that the resulting impulse response functions (IRF) can be quite misleading. In the current setting these restrictions imply that either a shock to returns doesn't affect contemporaneous flows (which contradicts the results of section III), or alternatively a shock to flows doesn't affect contemporaneous returns (which rules out any price pressure effects). The Choleski decomposition with $b_{21} = 0$ is equivalent to setting $\sigma_R^2 = b$ and $\sigma_f^2 = a - c^2 / b$ for the pure return and flow shocks respectively. Conversely, setting $b_{12} = 0$ is equivalent to setting $\sigma_f^2 = a$ and $\sigma_R^2 = b - c^2 / a$.

An alternative way to identify the VAR is to impose an additional restriction on the covariance matrix of the pure structural errors. This restricts the variance of pure flow shocks to lie within the extremes of $a - c^2 / b$ and a for σ_f^2. This is equivalent to setting $\sigma_f^2 = a - c^2 / b + k$ with $0 < k < c^2 / b$. Imposing this condition for a given k gives a quadratic equation that we can solve to give b_{12}. We can discard the negative solution due to the negative feedback trading result of section III that implies $b_{12} > 0$. The impulse response functions generated using this assumption are shown in Figure 8A & Figure 8B for settings of k=0 (equivalent to the standard Choleski approach with $b_{21} = 0$), and k= 0.01. The top panel of each figure shows the reaction of flows and returns to a one-standard deviation shock to flows. The bottom panel shows the reaction to a one-standard deviation shock to returns. A bootstrapping procedure is used to generate one standard error bands for the impulse response functions, following the method suggested by Hamilton (1994)[27].

The major difference between the two IRF's generated is the contemporaneous reaction of returns to flows. In the Choleski there is no immediate effect (due to the restriction). In the second IRF, returns react positively to a shock to returns.

[27] First the VAR is estimated using the original return and flow series, yielding an estimate of the reduced form phi parameter matrices. The residuals from this equation are used as an empirical distribution for the bootstrap. 1000 bootstrap samples are generated drawing from this distribution using these estimated phi parameter matrices. For each sample, a new path for returns and flows is generated and the phi matrices are re-estimated. Using these new phi estimates, the restriction on the variance of flows is imposed and the corresponding structural VAR is identified. This generates a new impulse response function for each bootstrap sample. Finally, the standard error of the impulse response function is estimated from the distribution of the 1000 draws of the impulse response function from the bootstrapping process.

We can economically interpret the IRF's as follow. For a pure positive return shock, small investors trade in a negative feedback fashion, leading to immediate negative flows. These flows dampen the original positive return shock, leading to short term under-reaction to new information. Flows are persistently negative for several weeks after the return shock, delaying the incorporation of new information into prices. Conversely for a positive flow shock, returns increase contemporaneously and in the following week consistent with short-term price pressure. The short-term price pressure is reversed in the subsequent week, despite the strong persistence of flows.

In summary, this section shows that there is a *positive* relationship between flows and short-term returns at the aggregate market level[28]. This result appears to be quite robust for full-service brokerage clients. At the market level flows predict one week ahead market returns, but short term returns exhibit reversals in the following week. This result is consistent with short-term price pressure induced by the aggregate actions of small investors moving into or out of the market.

B. Cross Sectional Flows

In this section the relationship between retail investor flows and returns at the cross-sectional level is examined. Net flows are also examined separately for full-service and internet brokerage clients. To estimate the relationship I estimate a cross-sectional reduced form VAR system with the following specification:

$$\begin{bmatrix} f_{it} \\ r_{it} \end{bmatrix} = \begin{bmatrix} \alpha_f \\ \alpha_r \end{bmatrix} + \begin{bmatrix} \phi_{11}(L) & \phi_{11}(L) \\ \phi_{11}(L) & \phi_{11}(L) \end{bmatrix} \begin{bmatrix} f_{it-1} \\ r_{it-1} \end{bmatrix} + \begin{bmatrix} e_{it}^{f} \\ e_{it}^{r} \end{bmatrix} \tag{6}$$

where $f_{i,t} =$ net flow for stock i in week t, $r_{i,t} =$ return for stock i in week t.

As in Vuolteenaho (2002), the phi parameters are assumed to be the constant across each stock to reduce the number of parameters to estimate. Parameters are estimated using equation-by-equation Fama-Macbeth OLS estimates on the panel of stocks. The VAR system is estimated with 12 lags in order to capture the full dependence of flows and returns on prior values. In the estimation, returns are cross-sectionally demeaned and standardised by the cross-sectional standard deviation to remove time-specific common shocks to mean returns and to reduce the effects of time-varying cross-sectional dispersion. In the robustness checks, various alternative excess return measures are also used.

Table 9 shows the parameter estimates for the reduced form VAR on cross sectional flows. Net flows are strongly persistent over time and negatively related to past returns, as shown in the earlier results. Returns also show the usual autocorrelation patterns identified in the literature, with short term (1-2 week) mean reversion (Conrad, Gultekin and Kaul (1997)) and medium term (3-12 week+) momentum (Jegadeesh and Titman (1993)). Additionally, the most novel finding is that retail investor flows *positively* and significantly predict returns over the following two weeks. This effect seems to have a permanent price impact, as there is no evidence of subsequent reversals in returns (in contrast to the aggregate market results). Splitting the net flows up into internet and full-service flows over the reduced 1996-2002 period shows that this positive predictability is again driven predominantly by the net flows of *full-service* brokerage clients.

Is this predictability result robust? To assess this issue, firstly I re-estimate the VAR system using excess returns calculated by matching each stock to 10 size benchmark portfolios. This doesn't effect the result, with the t-statistics on flow(t-1), flow(t-2), flow(t-3) in the return equation being 3.78 (vs 4.29), 2.83 (vs 3.59) and 2.09 (vs 1.85) respectively. Similarly using 25 (5x5) size & value sorted matching portfolios to calculate excess returns, or 27 (3x3x3) size, value, 6-month momentum matching portfolios do not alter the predictability result[29]. In each case past flows predict future short-term excess returns[30].

As an additional check I re-estimate the VAR parameters for full-service brokers for each individual year 1992-2002. The coefficient estimates for the sum of the flow (t-1) and flow (t-2) coefficients are shown in figure 9. For comparison, the coefficients for the other broker types are also included; institutional, mixed and internet. The results show that the coefficients for the full-service brokers are positive in *every* single yearly sub-period for the full-service brokers. This is not the case for the institutional or mixed brokers. The coefficients for the internet brokers are smaller than those of the full-service brokers. As a third robustness check, I estimate the equations for small and large stocks separately. Large stocks are defined as the largest 100 stocks in each weekly cross-section, while small stocks are all stocks in the index, outside the top 100. For brevity, I only report the coefficients in the return prediction equation. We can see from Table 10, that the basic return

---

[28] The relationship for longer-term returns is not examined in this paper. Kelly (1997) finds that individual investor participation predicts low aggregate market returns over the following one and three years.

[29] For 25 size value portfolios the t-stats are 4.01, 3.81 & 1.14 and for the 27 size, value, momentum portfolio the t-stats are 4.22, 3.84, 0.31 respectively. The results are not effected.

[30] Using unstandardised returns doesn't alter the result either. The t-statistics on the flow coefficients are 4.88, 2.83 & 2.09 respectively indicating the results are if anything slightly stronger.

patterns are similar for large and small stocks (short term mean reversion, medium term momentum). Additionally, small investor flows predict future weekly stock returns for *both* large and small stocks (albeit with a lag for smaller stocks). Again, full-service broker flows appear to be driving this effect.

An alternative approach to assess the robustness of these results is to construct portfolios using a methodology similar to that in Pan & Poteshman (2003). Each week I form a long and a short portfolio based on the net flows of the four different broker types. Stocks are allocated to the long portfolio if the broker type is a net buyer, and to the short portfolio if a net seller. The returns of the long and short portfolio are assessed over a window of t-5 to t+10 weeks for each of the 4 broker types (full-service, internet, mixed and institutional). Table 11 Panel A presents the returns of a long short portfolio based on this strategy around the weekly event date. We can see that the full-service broker zero investment strategy has a positive 21bp return in week t+1 (t-stat = 7.9), and a return of +38bp over the first four weeks. Stocks in this portfolio have been poor previous performers (shown by the strongly negative return difference from t-5 to t). Internet brokers are also net buyers of poorly performing stocks, however they exhibit no positive performance post the portfolio formation date. Institutional brokers are net buyers of strongly performing stocks, but also have no positive subsequent short-term outperformance. Panel B presents the long short returns from an alternative strategy, where stocks are assigned to the long portfolio only if the net flow ratio is greater than 0.25, and the short portfolio if less than –0.25. The economic finding is unchanged, the full-service broker strategy has a positive return of +30bp in week t+1, and the internet and institutional brokerage strategies show no excess returns.

These results suggest that the flows of full-service brokers are useful in predicting future short-term cross-sectional stock returns. These results complement those of Pan & Poteshman (2003) who show in the options market that the trades of full-service brokerage clients contain information about future short-term stock returns. The difference in forecasting ability of the internet and full-service brokerage clients is consistent with the latter paying higher brokerage commissions in return for valuable private information. Choe, Kho & Stulz (2000) also find that domestic individuals have valuable short-term private information in the Korean market[31]. Alternatively Griffin et al (2003) do not find evidence that individual order imbalances predict future returns in 9-months of Nasdaq data, however they do not separate internet from full-service investors.

---

[31] Froot, O'Connell & Seasholes (2001) examine international investor flows and find evidence that flows predict future market returns in emerging equity markets. They also employ a bivariate VAR methodology in their analysis.

An interesting question is why full-service broker flows predict future returns? Given no evidence of subsequent reversals, temporary price pressure seems unlikely. One possible explanation is that these investors have valuable private information. Alternatively, individual investors may be acting as liquidity providers to institutional brokers, allowing them to unload inventories of stock acquired through principal trades. Institutional brokers cross much of their stock internally via upstairs trading, and might be willing to pay on average a small premium to small investors to take stock that cannot be disposed of in this system. Future research will explore these possible explanations in more detail.

Finally I estimate the impulse response function for net flows and stock returns using the variance restriction outlined in section V.A. The estimated impulse response functions are shown in Figure 10. A positive flow shock leads to a persistent flow reaction in following weeks. Additionally short term returns are increased at week 0, and slightly at weeks 1 and 2 consistent with the reduced form results. A positive shock to returns leads to an immediate and persistent negative shock to flows, consistent with negative feedback trading by small investors. The positive return shock also generates well-documented patterns of one-week mean reversion and medium term price momentum.

In summary section V shows there is a strong relationship between aggregated small investor flows and short-term returns. In contrast to models that assume individual investors are irrational, this section shows that the flows of these investors actually positively predict short-term returns at both the aggregate and cross-sectional levels. Further examination reveals that full-service brokerage clients predominantly drive this result.

VI. Conclusion

This paper uses Australian data over the period 1991 to 2002 to examine whether individual investor's trades contain systematic patterns that survive after aggregation. Aggregate individual trades reveal a strong negative relationship between net flows and returns indicating that individual investors are *negative feedback traders* at a weekly frequency. This is true for both aggregate market level flows (flows into and out of the market as a whole) and cross-sectional flows. The data also reveals strong persistence in flows, which is especially prevalent at the cross-sectional level. Small investors are net buyers of stocks with high recent volatility. Aggregate small investor trading activity is high in the first half of January and in the second half of June (before tax-year end).

Examining net flows at the cross-sectional level individually for 56 retail brokers reveals that net flows across independent firms are significantly positively related. The average correlation between weekly cross-sections of net flows for the typical internet brokerage firm pair is 0.23 and for full-service is 0.16. For top 50 stocks only these correlations increase to 0.44 and 0.23 respectively. Additionally *every* single retail broker pair has a significantly positive correlation coefficient. These results confirm that the actions of individual investors are indeed highly pervasive across a wide range of unrelated brokerage firms.

The final section examines the relationship between small investor flows and future returns. At the market level, I find that the net flows by full-service brokerage clients predict future short-term aggregate market returns. This finding is consistent with short-term price pressure given evidence of subsequent reversals. At the cross-sectional level I find that full-service clients net trades positively predict returns for a period of up to 2 weeks with no evidence of subsequent reversals. This finding is consistent with small investors using valuable private information, or acting as liquidity providers allowing institutional brokers to unload inventory positions.

In summary, the main finding of this paper is that individual investor trades *are* highly systematic across a large number of brokerage firms. This implies individual trading patterns aggregate strongly and have the potential to influence market dynamics. However the conjecture that individual investors are irrational may not be valid since the flows of full-service investors *positively* predict future short-term market and cross-sectional returns.

Future research could further examine the interaction between institutional and retail investors to distinguish liquidity provision from private information explanations. Studies using individual investors trades and other sources of disaggregated order flow have the potential to reveal much about the real-life behaviour of different types of investors and their impact on financial markets.

References

Aitken, M, Kua, A, Brown, P, Walter, T & Izan, H (1995), An Intraday Analysis of the Probability of Trading on the ASX at the Asking Price, *Australian Journal of Management,* 20(2), 115-154.

Baker, M & Stein, J (2002), Market Liquidity as a Sentiment Indicator, *NBER Working Paper*, w8816.

Barber, B and Odean, T (2000), Trading Is Hazardous To Your Wealth: The Common Stock Investment Performance of Individual Investors, *Journal of Finance,* 55(2) 773-806.

Barber, B and Odean, T (2002), All That Glitters: The Effect of Attention and News on the Buying Behaviour of Individual and Institutional Investors, *Working Paper*, University of California.

Barber, B, Odean, T and Zhu, N (2003), Systematic Noise, *Working Paper*, University of California.

Barberis, N & Thaler, R (2001), A Survey of Behavioural Finance, *Working Paper, University of Chicago*.

Benartzi, S (2001), Excessive Extrapolation and the Allocation of 401(k) Accounts to Company Stock?, *Journal of Finance*, 56(5), 1747-1764.

Choe, H, Kho, B and Stulz, R (2000), Do domestic investors have more valuable information about individual stocks than foreign investors, *NBER Working Paper.*

Cohen, R (1999), Asset Allocation Decisions of Individuals and Institutions, *Working Paper, Harvard University.*

Conrad, J, Gultekin, M and Kaul, G (1997), The Profitability of Short-Term Contrarian Strategies: Implications for Market Efficiency, *Journal of Business and Economic Statistics*, 15(3), 379-386.

De Long, J, Shleifer, A, Summers, L and Waldmann, R (1990), Noise trader risk in financial markets, *Journal of Political Economy*, 98, 703-738.

Dhar, R and Kumar, A (2001), A Non-Random Walk Down the Main Street: Impact of Price Trends on Trading Decisions of Individual Investors, *Working Paper Yale*.

Edelen, R and Warner, J (2001), Aggregate Price Effects of Institutional Trading: A Study of Mutual fund Flow and Market Returns, *Journal of Financial Economics*, 59, 195-220.

Enders, W (1995), *Applied Econometric Time Series*, Wiley: New York.

Epps, T and Epps, M (1976), The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis, *Econometrica*, 44, 305-321.

Fong, K, Madhavan, A & Swan, P (2001), Why do Markets Fragment? A Panel-Data Analysis of Off-Exchange Trading, *Working Paper*, University of Sydney.

Froot, K, O'Connell, P & Seasholes, M (2001), The portfolio flows of international investors, *Journal of Financial Economics*, 59, 151-193.

Gallant, R, Rossi, P and Tauchen, G (1992), Stock Prices and Volume, *Review of Financial Studies*, 5(2), 199-242.

Griffin, J, Harris, J & Topaloglu, S (2003), The Dynamics of Institutional and Individual Trading, *Journal of Finance*, forthcoming.

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

Grinblatt, M and Keloharju, M (2001), What Makes Investors Trade, *Journal of Finance,* 56(2) 589-616.

Hamilton, J (1994), *Time Series Analysis*, Princeton, New Jersey.

Hong, H & Stein, J (1999), A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets, *Journal of Finance*, 54(6), 2143-2184.

Jegadeesh, N and Titman, S (1993), Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency, *Journal of Finance*, 48(1), 65-91.

Kahneman, D and Tversky, A (1982), *Judgement under Uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge.

Karpoff, (1987), The Relation between Price Changes and Trading Volume: A Survey, *Journal of Financial & Quantitative Analysis*, 22(1), 109-126

Kelly, M (1997), Do Noise Traders Influence Stock Prices?, *Journal of Money, Credit and Banking*, 29(3), 351-363

Lee, C, Shleifer, A and Thaler, R (1991), Investor sentiment and the closed-end fund puzzle, *Journal of Finance,* 46, 75-110.

Lo, A and Wang, J (2000), Trading Volume: Definitions, Data Analysis, and Implications of Portfolio Theory, *Review of Financial Studies*, 13(2), 257-300.

Newey, W and West, K (1987), A Simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55, 703-708.

Odean, T (1998), Are investors reluctant to realise their losses?, *Journal of Finance*, 53(5), 1775-1798.

Pan, J and Poteshman, A (2003), The Information in Option Volume for Stock Prices, *MIT Working Paper* 4275-03.

Poteshman, A and Serbin, V (2003), Clearly Irrational Financial Market Behavior: Evidence from the Early Exercise of Exchange Traded Stock Options, *Journal of Finance*, 58 (1), 37-70.

Ranguelova, E (2001), Disposition Effect and Firm Size: New Evidence on Individual Investor Trading Activity, *Working Paper*, Harvard.

Ritter, J (1988), The Buying and Selling Behaviour of Individual Investors at the Turn of the Year, *Journal of Finance*, 43(3), 701-717.

Shapira, Z and Venezia, I (2000), Patterns of behaviour of professionally managed and independent investors, *Journal of Banking and Finance*, 25(8), 1573-1587.

Shefrin, H and Statman, M (1985), The disposition to sell winners too early and ride losers too long: Theory and evidence, *Journal of Finance*, 40, 777-790.

Shiller, R (1998), Human Behaviour and the Efficiency of the Financial System, *Cowles Foundation Working Paper*.

Shleifer, A (2000), *Inefficient Markets*, Oxford: Oxford University Press.

Sims, C (1980), Macroeconomics and Reality, *Econometrica*, 48, 1-49.

Vuolteenaho, T (2002), What Drives Firm Level Stock Returns, *Journal of Finance*, 57(1), 233-264.

Warther, V (1995), Aggregate Mutual Fund Flows and Security Returns, *Journal of Financial Economics*, 39, 209-235.

APPENDIX I: DETRENDING OF MARKET LEVEL FLOWS

Market level buy, sell and total flows exhibit substantial non-stationarity and positive skewness. To model these variables we need to transform them to achieve stationarity. Lo & Wang (2000) discuss various transformation methods that have previously been used for transforming volume data including removing linear or log linear time trends, taking differences and using non-parametric kernel regressions.

Several obvious factors affect the low frequency dynamics of the flows in this sample including the growing market share of the brokers in the sample (see Figure 1A), and the growth in overall market turnover and market capitalisation over the period. In order to determine the appropriate de-trending methodology, the log(buy), log(sell) and log(total) flows were first evaluated for the presence of a stochastic trend (unit root). An augmented Dickey-Fuller test with time trend rejects the presence of a stochastic trend for all three series at the 1% significance level[32]. To preserve the information in the levels (which is destroyed by taking differences), the data is de-trended using a deterministic polynomial function of order five[33] and the residuals are standardised by their mean and standard deviation. The fitted polynomial fitted trend looks similar to a linear time trend with some curvature to capture the rapid growth of the small brokerage volume in the 1997-2000 period and subsequent plateauing of this series. Detrended series are also standardized to have mean of zero and standard deviation of one.

To assess the robustness of the buy, sell and total flow results I repeat the analysis in Table 2 for buy, sell and total flows using firstly raw flows and secondly flows detrended using a Hodrick-Prescott filter. This produces the same significant return, flow, volatility and calendar effects. Repeating the analysis using log-differenced flows also produces the same effects for past returns, volatility and calendar variables, but gives negative auto-correlations for the past flow variables due to the over-differencing induced by the transformation.

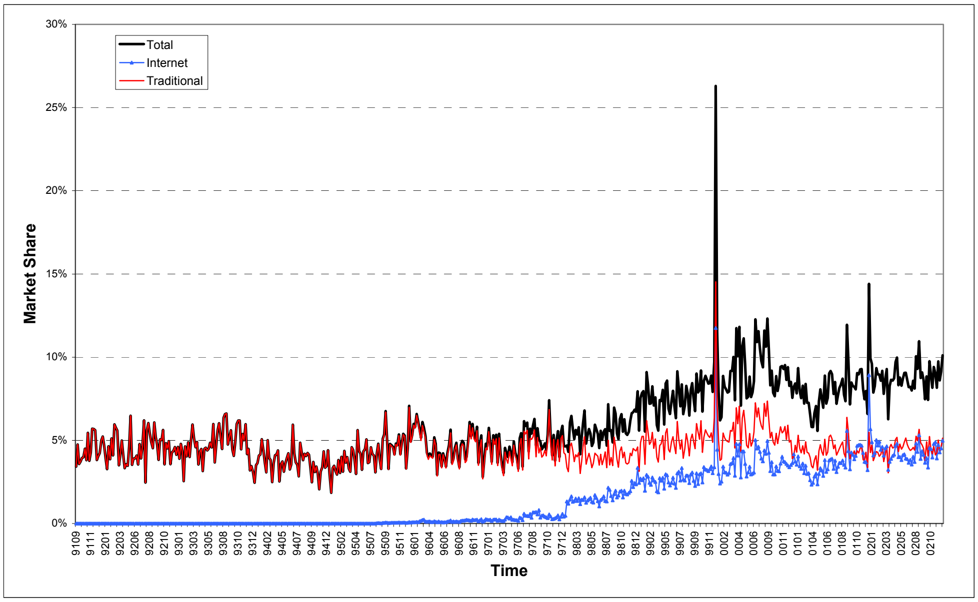
---

[32] The lag length used in the unit root test was one, which was determined using the AIC criterion. Examining the autocorrelation function for first differences of the buy, sell and total flow data shows no significant lags other than the first.

[33] The polynomial order was determined using the Schwartz information criterion. For polynomials of order > 5 the Schwartz criterion is reasonably flat (the AIC continues to fall). However, for too high an order polynomial, we are no longer removing low frequency information from the series, but start to remove the high frequency information, which defeats the purpose of the de-trending. The fifth order polynomial has enough degrees of freedom to fit the low frequency dynamics, without removing the high frequency information.

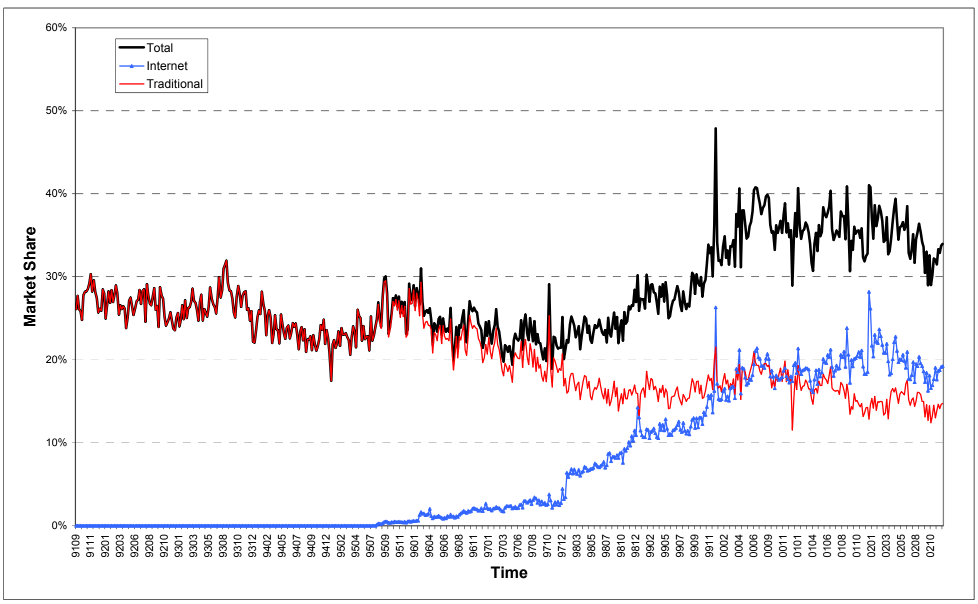
APPENDIX II: FIGURES

FIGURE 1A: AGGREGATE MARKET SHARE OF SAMPLED BROKERS OVER TIME (BY $ VALUE)



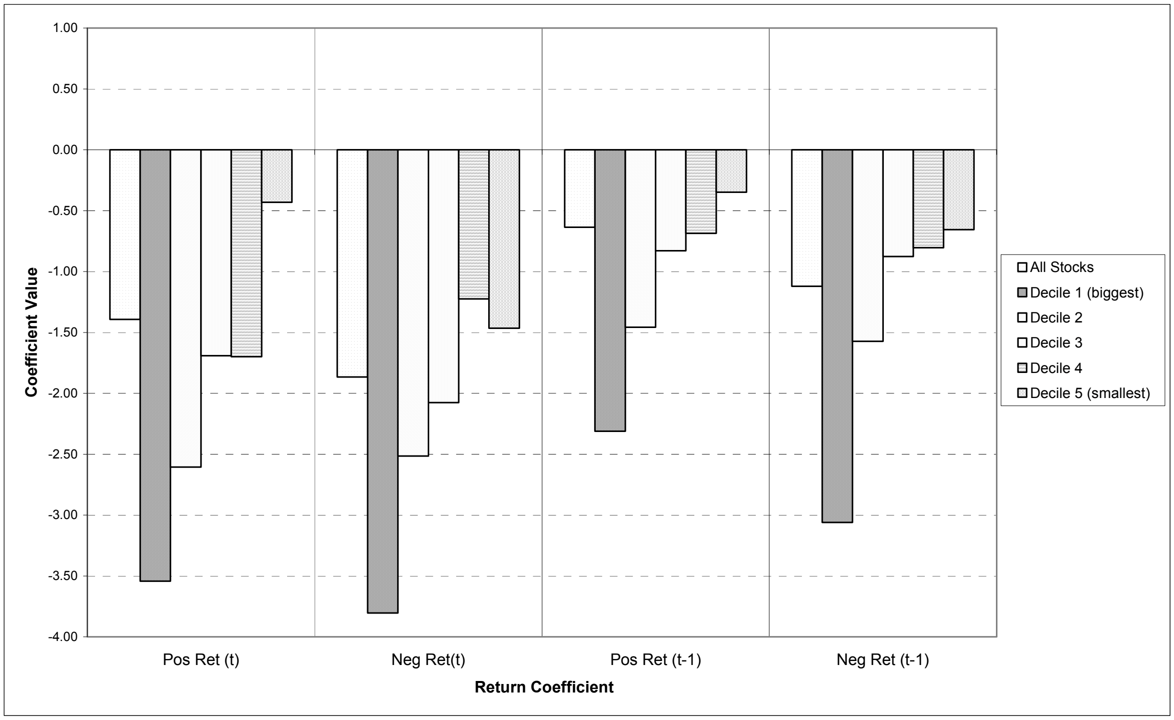
This figure shows the market share of the brokers in the sample from September 1991 to December 2002. Market share is calculated as the value of trades done through these brokers relative to the value of total market turnover.

FIGURE 1B: AGGREGATE MARKET SHARE OF SAMPLED BROKERS OVER TIME (BY # TRADES)



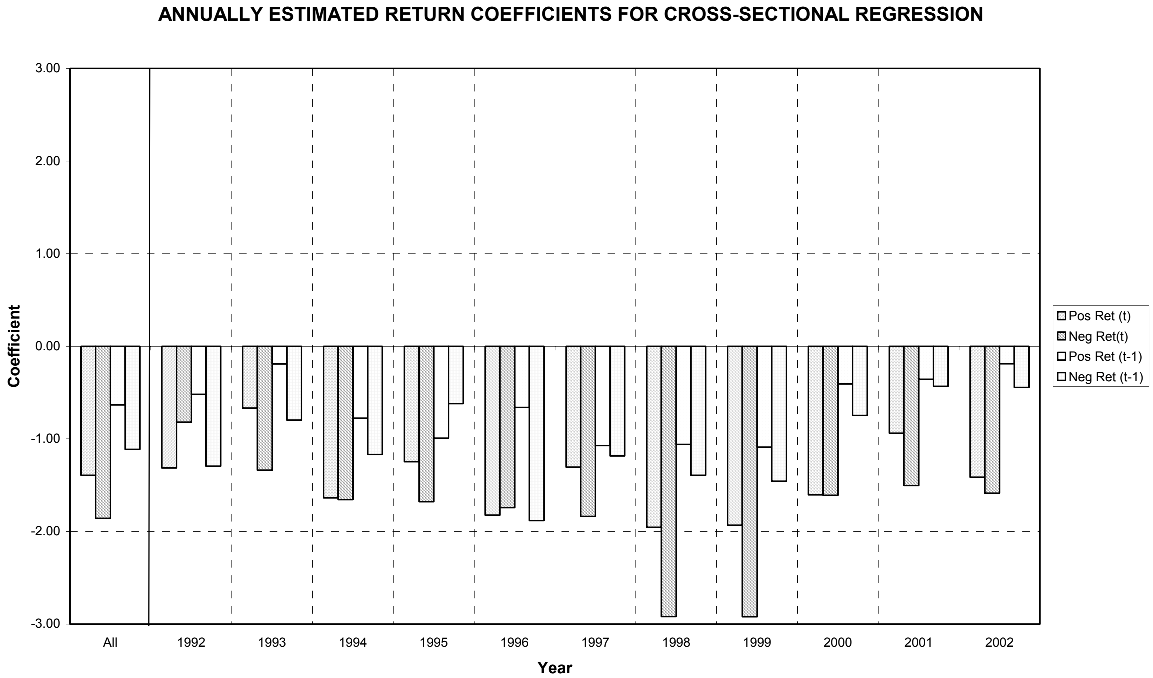
This figure shows the market share by number of trades of the brokers in the sample from September 1991 to December 2002. Market share is calculated as the number of trades done through these brokers relative to the total number of trades.

FIGURE 2: RETURN COEFFICIENTS ESTIMATED FOR DIFFERENT SIZE QUINTILES



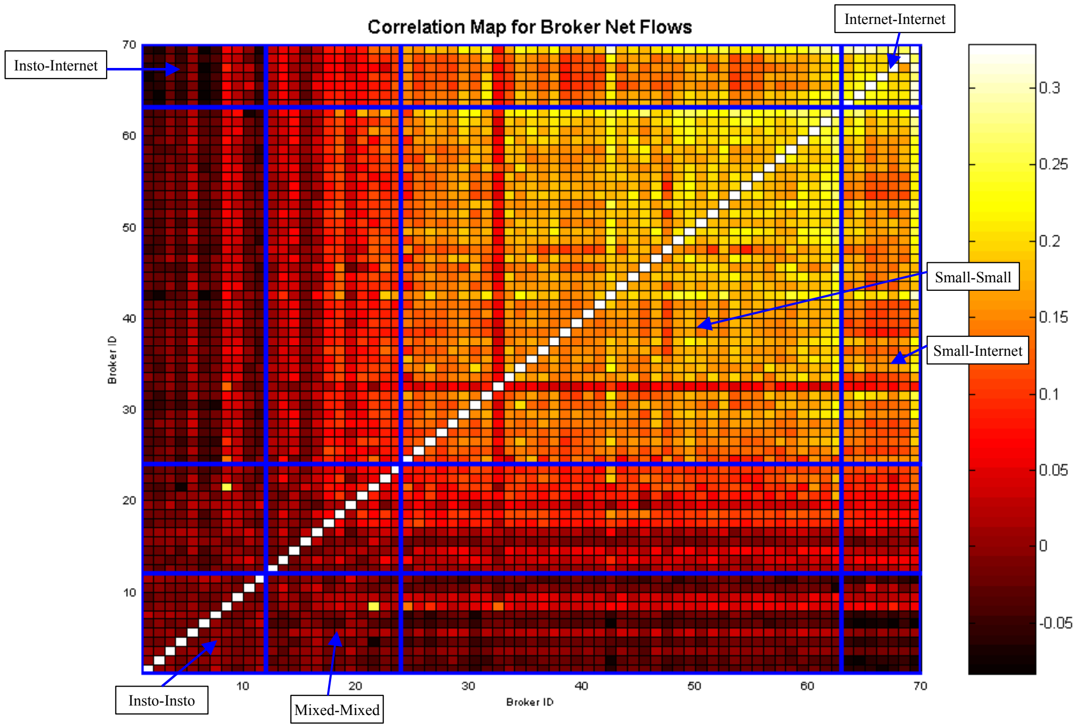
This figure shows the return coefficients for the four asymmetric regressors corresponding to the regression in Table 4. The coefficients are estimated firstly on all stocks (leftmost bar in each group of six), then on quintiles of stocks starting with the largest stock group second from the left, and moving to the smallest quintile of stocks (on the rightmost of each group of six). This figure demonstrates that the regression coefficients are of consistent sign for both large and small stocks, and are stronger for larger stocks.

FIGURE 3: RETURN COEFFICIENTS FOR DIFFERENT SUB-PERIODS



This figure shows the return coefficients for the four asymmetric regressors corresponding to the regression in Table 4. The regressions are estimated separately in each year and the coefficients for the PosRet(t), NegRet(t), PosRet(t-1) and PosRet(t-2) are reported above. This shows that in every subperiod these coefficients are significantly negative.

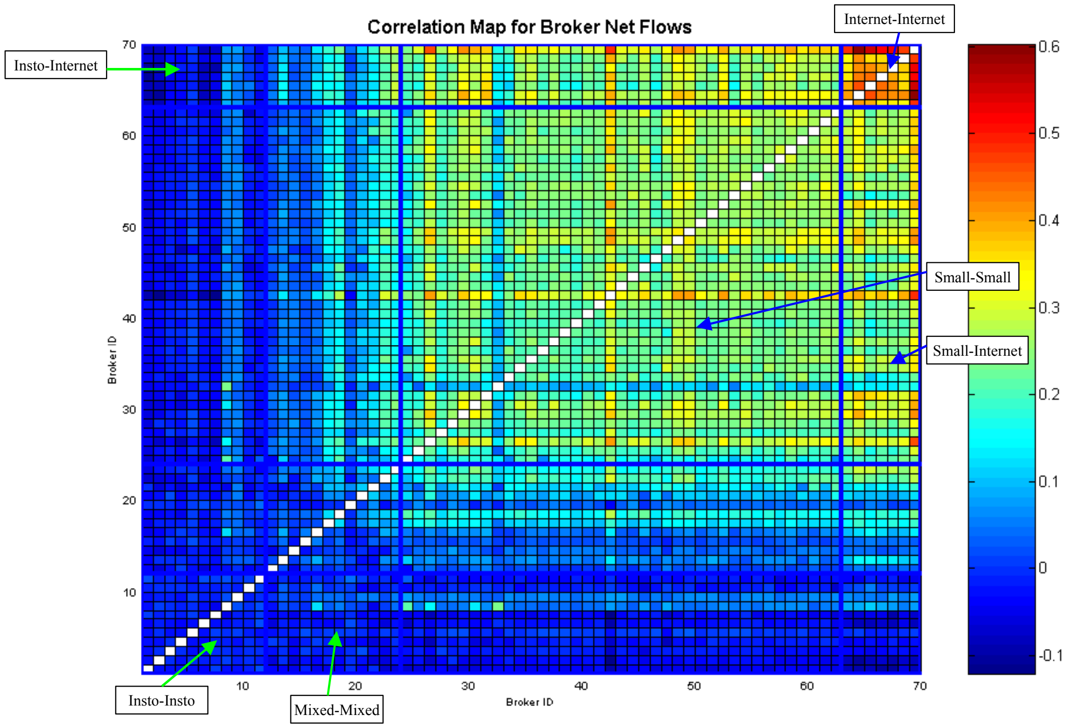
FIGURE 4: CORRELATION MAP FOR BROKER NET FLOWS – ALL STOCKS



This figure shows the average weekly cross-sectional correlation between the net flows from 70 separate brokerage firms for all index stocks. Brokers are split into 11 institutional brokers (ID: 1-11), 12 mixed brokers (ID: 12-23), 39 full-service small brokers (ID: 24-62) and 8 internet brokers (ID: 63-70). Within each block, brokers are sorted in descending order of average trade size, so brokers with smaller clients will appear to the right of each block.

The period of the analysis is September 1991 to December 2002. Dark colours represent areas of low or negative correlation; lighter colours indicate higher correlations as shown in the colourbar on the right of the plot. The labels indicate the type of correlation in each sub-block eg 'Small-Internet' refers to the correlation of the 9 internet brokers with the 39 full-service small brokers.

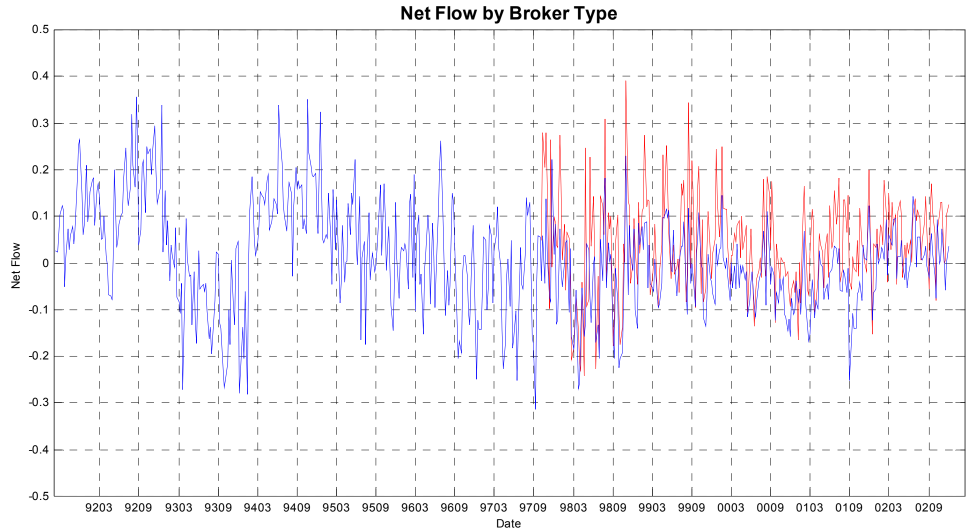
FIGURE 5: CORRELATION MAP FOR BROKER NET FLOWS – TOP 50 STOCKS



This figure shows the average weekly cross-sectional correlation between the net flows from 70 separate brokerage firms for the top 50 index stocks only. Brokers are split into 11 institutional brokers (ID: 1-11), 12 mixed brokers (ID: 12-23), 39 full-service small brokers (ID: 24-62) and 8 internet brokers (ID: 63-70). Within each block, brokers are sorted in descending order of average trade size, so brokers with smaller clients will appear to the right of each block.

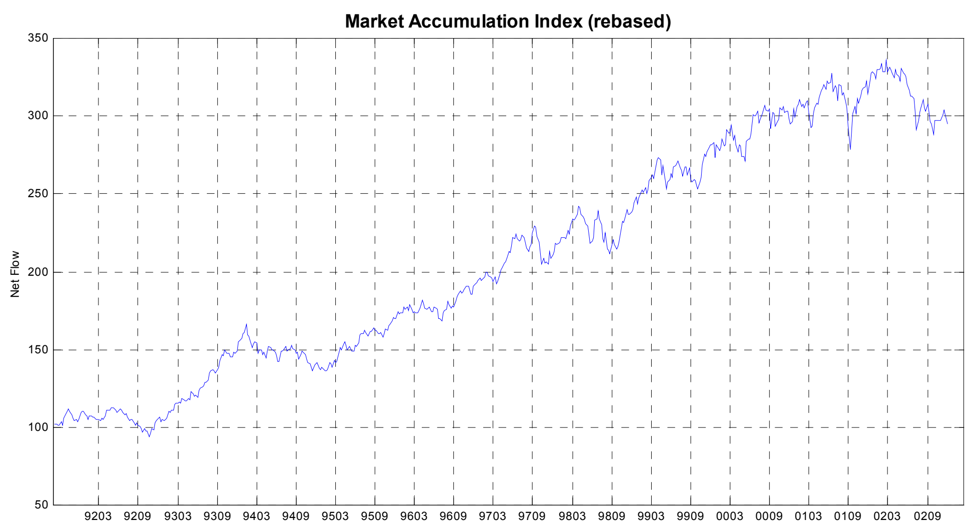
The period of the analysis is September 1991 to December 2002. Dark colours represent areas of low or negative correlation; lighter colours indicate higher correlations as shown in the colourbar on the right of the plot. The labels indicate the type of correlation in each sub-block eg 'Small-Internet' refers to the correlation of the 9 internet brokers with the 39 full-service small brokers.

FIGURE 6A: MARKET LEVEL NET FLOWS FOR FULL-SERVICE & INTERNET BROKERS



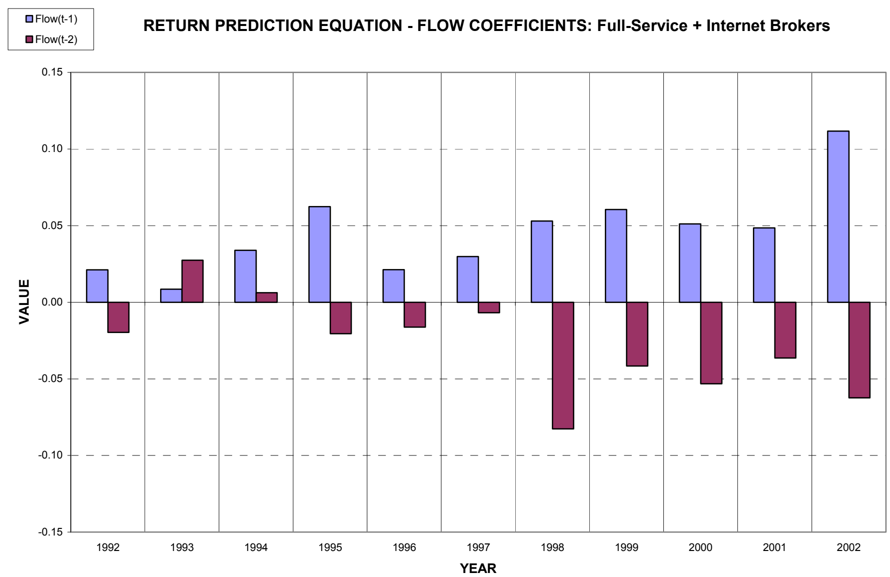
This figure shows aggregate market level net flows for internet and full-service brokers over the sample period from September 1991 to December 2002.

FIGURE 6B: MARKET ACCUMULATION INDEX (REBASED)



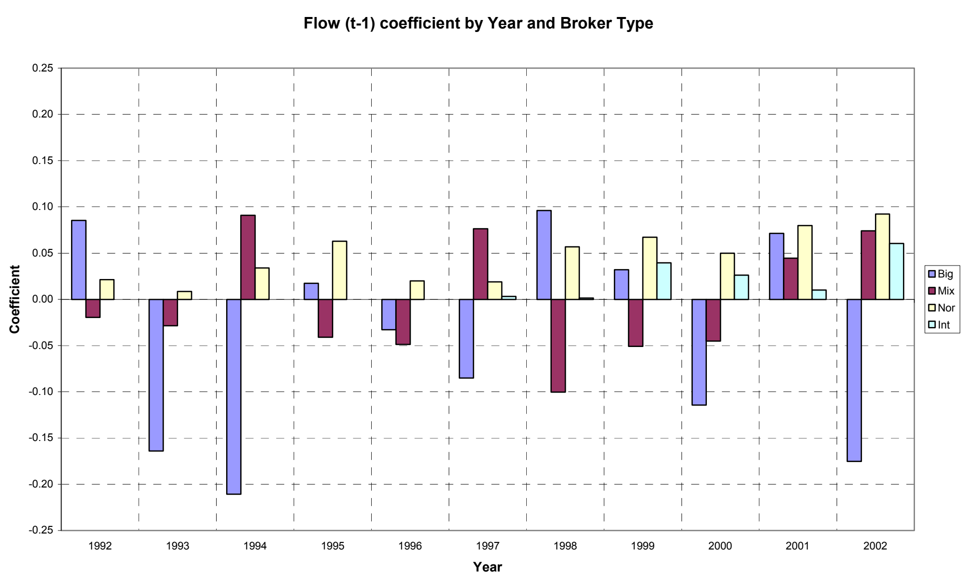
This figure shows the market accumulation index over the sample period from September 1991 to December 2002.

FIGURE 7A: FLOW COEFFICIENTS IN RETURN FORECASTING EQUATION: YEAR BY YEAR



This figure shows the coefficients for flow(t-1) in blue and flow(t-2) in red for the return prediction regressions of Table 8. The coefficients are estimated year-by-year using the aggregate market net flows of all brokers. We can see that the flow(t-1) coefficient is positive in every single year in the sample.

FIGURE 7B: FLOW COEFFICIENTS BY BROKER TYPE: YEAR BY YEAR



This figure shows the coefficients for flow(t-1) in the return prediction regression of Table 8 by broker type for each separate year. The coefficients are estimated year-by-year using the aggregate market net flows of all brokers. The flow (t-1) coefficient is positive in every year for full-service & internet brokers, unlike for other broker types.

Figure 8A: Impulse Response Functions for Market Flows and Returns, K=0.0

The upper two panels in this figure show the response of flows and returns to a +1 standard deviation shock in returns. The lower two panels show the response to a +1 standard deviation shock in returns. One standard deviation error bands are calculated using the bootstrap method described in the paper. For this figure, returns and flows are expressed in standard deviation units. Panel 8A shows the results for Choleski decomposition.

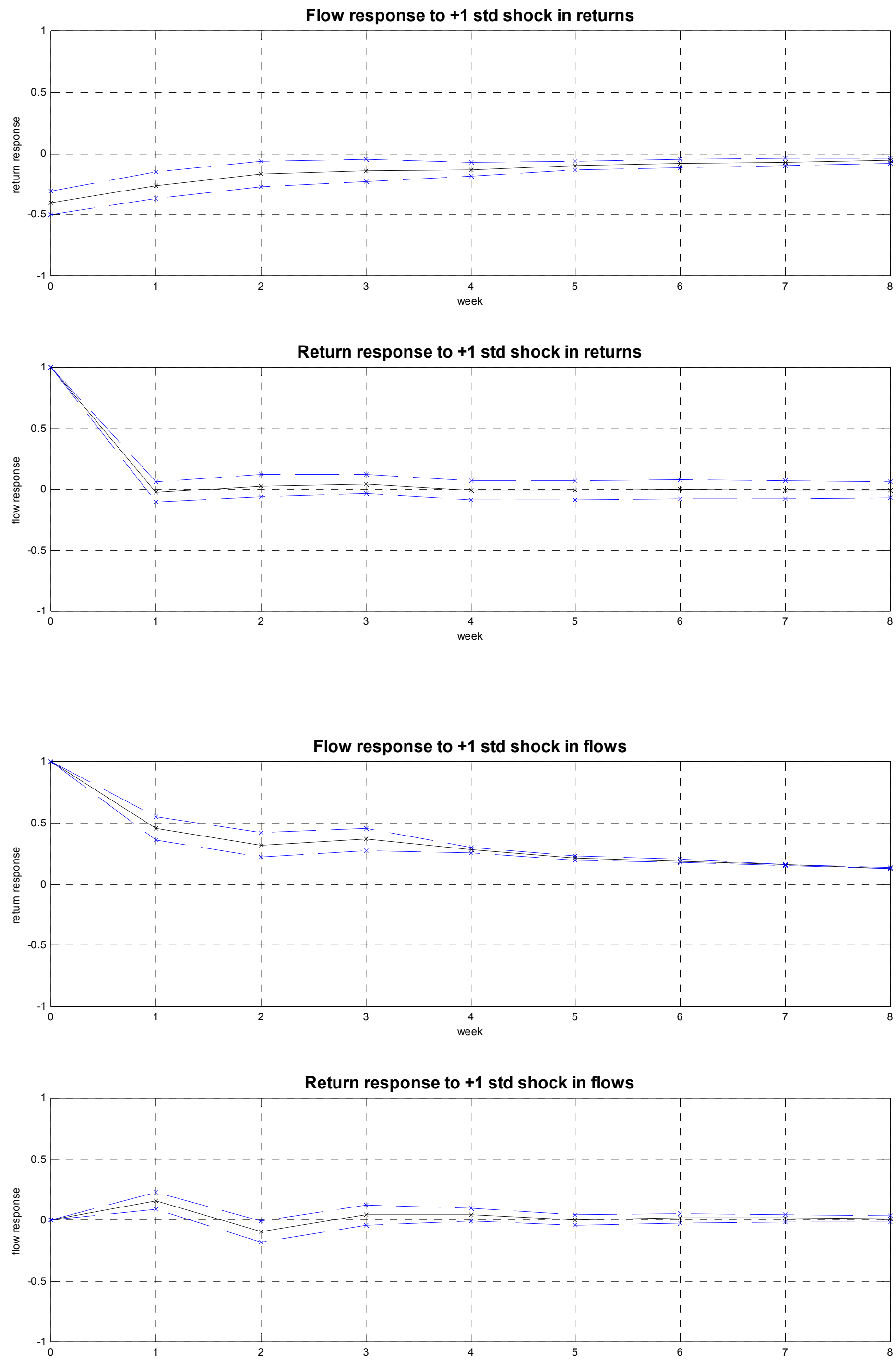


FIGURE 8B: IMPULSE RESPONSE FUNCTIONS FOR MARKET FLOWS AND RETURNS, K=0.01

The upper two panels in this figure show the response of flows and returns to a +1 standard deviation shock in returns. The lower two panels show the response to a +1 standard deviation shock in returns. One standard deviation error bands are calculated using the bootstrap method described in the paper. For this figure, returns and flows are expressed in standard deviation units. Figure 8B shows the results for the covariance restriction identification.

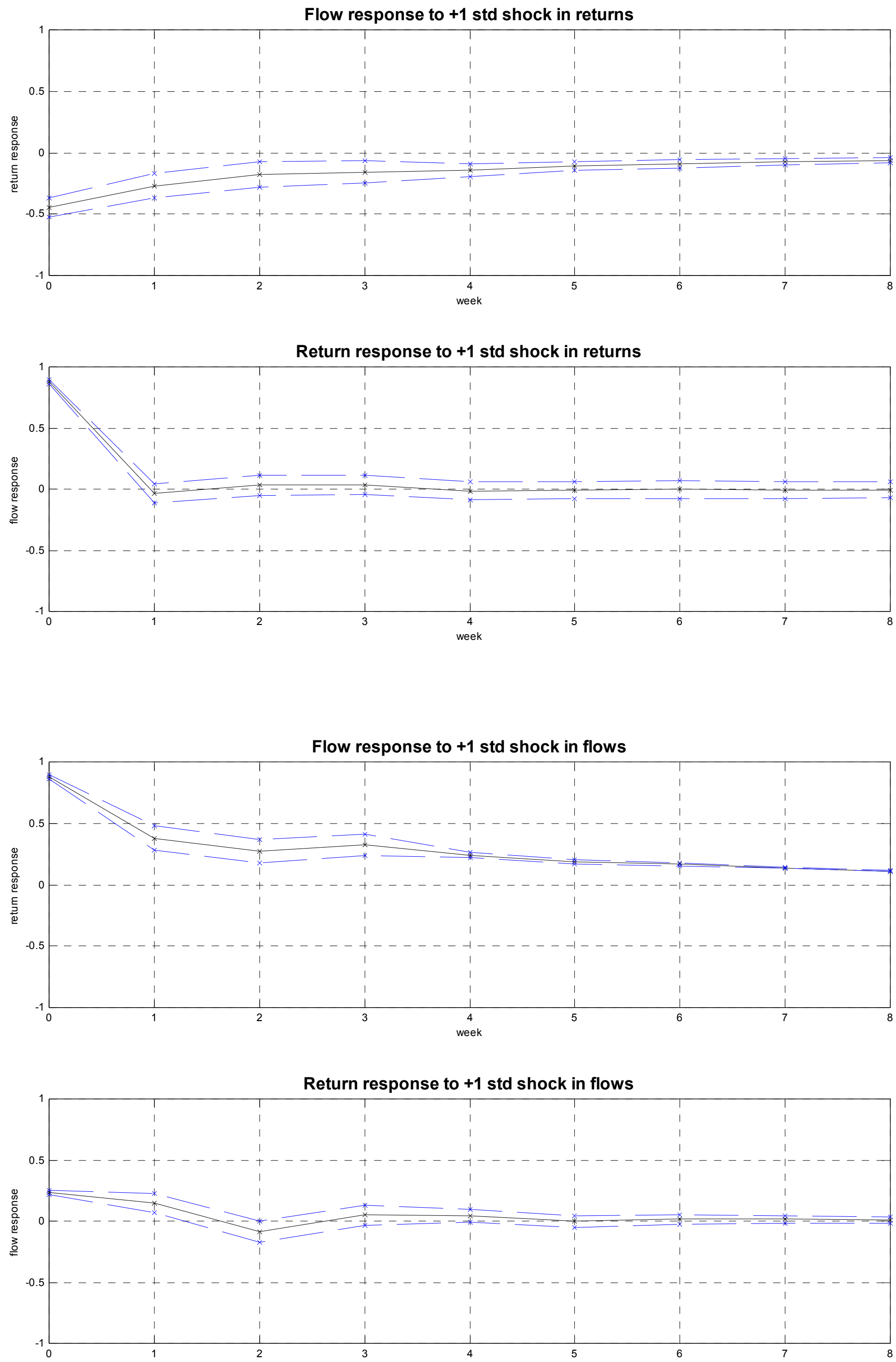


FIGURE 9: FLOW COEFFICIENTS IN RETURN FORECASTING EQUATION: YEAR BY YEAR

This figure shows the sum of the coefficients on flow(t-1) and flow(t-2) for the cross-sectional return(t) prediction regressions of Table 9. The coefficients are estimated year-by-year separately for each broker type. The sum of the coefficients is positive in every year for the full-service retail brokerage firms.

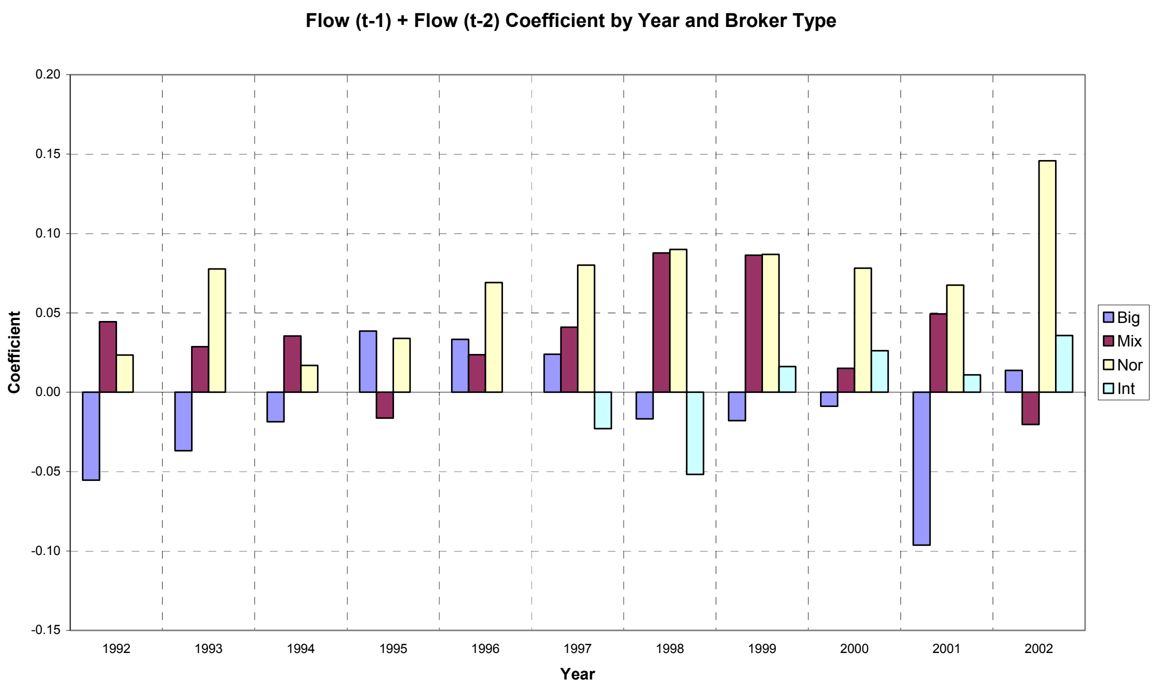
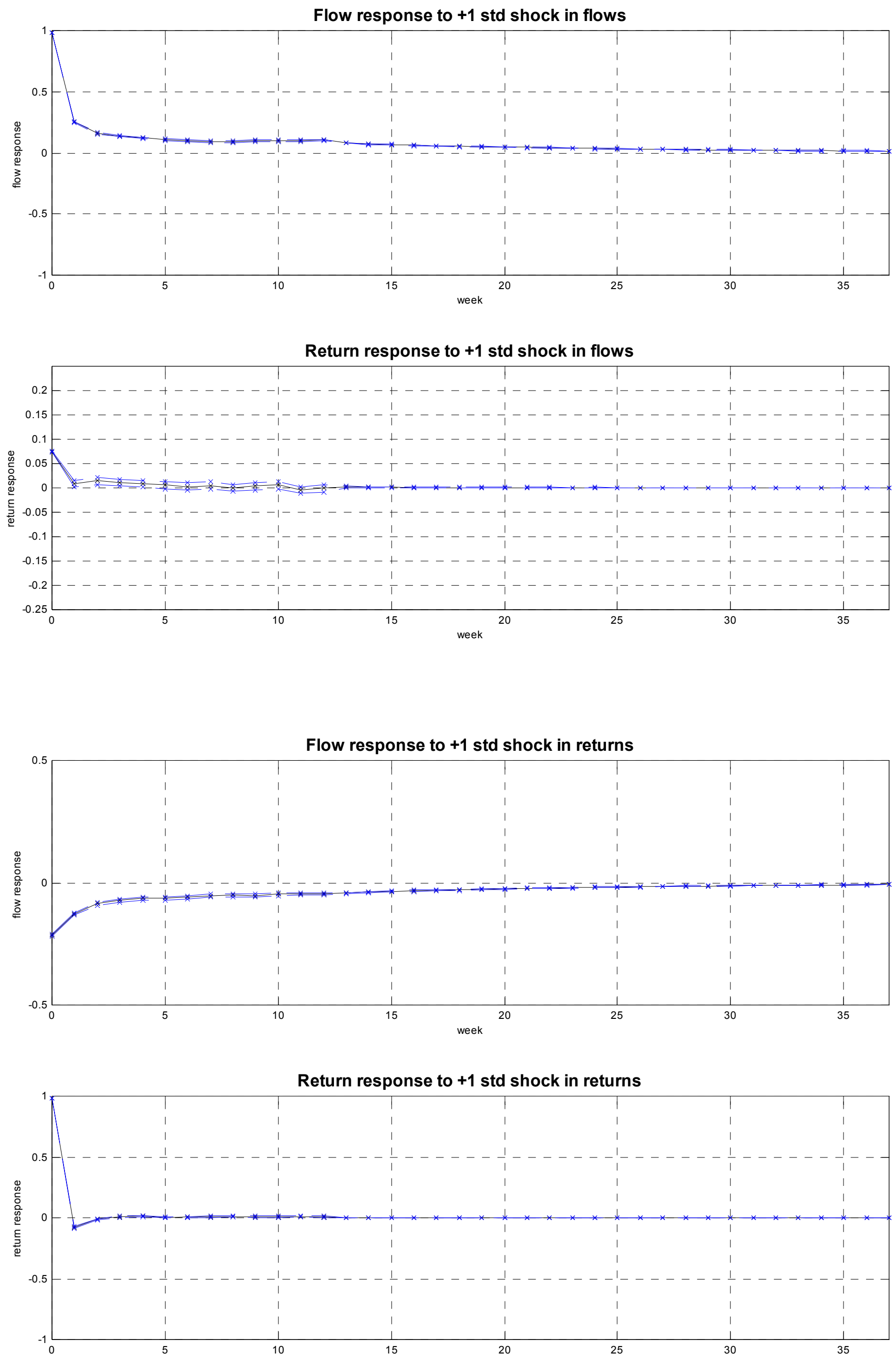


FIGURE 10: IMPULSE RESPONSE FUNCTIONS FOR CROSS-SECTIONAL FLOWS AND RETURNS

The upper two panels in this figure show the response of returns and flows to a +1 standard deviation shock in returns. The lower two panels show the response to a +1 standard deviation shock in returns. Both flows and returns are expressed in standard deviation units.



Appendix III: Tables

Table 1A: Summary Description of Aggregate Data (in $A)

This table presents summary statistics on the data sample. Results are split into internet brokerage investors and traditional full-service brokerage investors. The sample period is September 1991 to December 2002.

	Broker Type		
Variable	**Internet**	**Traditional**	**All**
Total # Buys	9,305,566	12,351,483	21,657,049
Total # Sells	7,723,715	12,509,032	20,232,747
Total Buy Value ($M)	67,271	123,839	191,110
Total Sell Value ($M)	61,152	126,861	188,012
Average Buy Trade	7,229	10,026	8,824
Average Sell Trade	7,917	10,142	9,292
# Brokers	9	47	56
Stock-Weeks	597,985	2,664,365	3,262,350

Table 1B: Market Share of Retail Brokers By Size Group

This table shows market share by value for the retail brokers for the average stock in each size portfolio. The left column shows the average market share for the full sample period, the right three columns for the final three years of the sample split into traditional full-service and internet brokerage firms.

Market Share by Size Portfolio		Sep 1991-Dec 2002 All 56 retail brokers	Jan 2000-Dec 2002 All 56 retail brokers	Jan 2000-Dec 2002 47 traditional brokers	Jan 2000-Dec 2002 9 internet brokers
Size Portfolio*	1-20	4.9%	6.2%	2.8%	3.5%
	21-50	7.2%	9.2%	3.9%	5.3%
	51-100	12.7%	13.9%	5.6%	8.3%
	101-200	25.4%	34.5%	14.7%	19.8%
	200+	41.9%	57.4%	26.7%	30.7%
	All	24.2%	33.2%	14.7%	18.5%

Table 1C: Descriptive Market Statistics for 2001 (or most recent available)

Country	# Individual Shareholders (m)	Participation Rate	% Market Owned by Retail Investors	Ownership by Foreign Investors	Turnover Velocity	# Listed Companies	Market Capitalisation (US$bn)
Australia	5.7	41%	24%	24%	67%	1410	375
Finland	0.8	18%	12%	51%	99%	155	190
France	5.6	13%	17%	14%	138%	1132	1844
Germany	na	na	17%	14%	118%	983	1072
Hong Kong	1.0	17%	na	na	44%	867	506
Ireland	0.4	13%	na	na	24%	87	75
Japan	32.1	30%	18%	19%	60%	2141	2294
UK	11.5	24%	16%	33%	84%	2332	2165
US	33.8	26%	42%	11%	141%	7069	13827

TABLE 2: AGGREGATE WEEKLY FLOW REGRESSIONS: NET FLOWS, BUY FLOWS, SELL FLOWS & TOTAL FLOWS

This table analyses aggregated weekly net flows, buy flows, sell flows and total flows. Aggregated weekly flows are calculated by summing buys and sells for all individual stocks across all brokers in the sample each week. Buy, sell and total flows are calculated as de-trended raw flows. Net flows are calculated as the (buyflow-sellflow)/(buyflow+sellflow). MktRet is the weekly market return and volatility is the weekly return volatility calculated using the daily market returns. T-stats are calculated using the Newey-West method with four lags. The period of the regression is September 1991 to December 2002 and the frequency is weekly.

Variable	Net Flows				Buy Flows				Sell Flows				Total Flows				Variable
	Full Model		Reduced model		Full Model		Reduced model		Full Model		Reduced model		Full Model		Reduced model		
	Coef.	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	
p-value	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000		p-value
adjR2	57.4%		58.0%		56.6%		56.9%		63.2%		63.7%		60.9%		61.3%		adjR2
AIC																	AIC
Constant	0.00	0.31	0.01	1.75	-0.13	-0.90	-0.07	-0.78	-0.06	-0.42	-0.11	-1.42	-0.09	-0.57	-0.13	-1.12	Constant
MktRet (t)	-2.73	-12.01	-2.71	-11.36	-2.85	-1.75	-2.92	-1.80	14.82	8.42	14.71	8.33	6.74	4.10	6.82	4.16	MktRet (t)
MktRet (t-1)	-0.77	-3.22	-0.58	-2.79	1.28	0.76			6.41	3.30	6.16	3.30	4.19	2.42	4.49	2.56	MktRet (t-1)
MktRet (t-2)	-0.30	-1.28			4.40	2.45	4.09	2.35	5.91	3.51	5.47	3.21	5.35	3.22	5.69	3.44	MktRet (t-2)
MktRet (t-3)	0.08	0.38			3.37	2.00	3.70	2.24	3.33	2.01	3.37	2.04	3.54	2.26	3.85	2.59	MktRet (t-3)
MktRet (t-4)	-0.04	-0.20			3.09	1.91	3.82	2.39	1.49	1.01			2.42	1.68	2.91	2.09	MktRet (t-4)
MktRet (t-5:t-8)	-0.15	-1.41	-0.15	-1.62	1.11	1.25	1.72	2.13	2.25	2.45	2.11	2.73	1.70	1.88	2.01	2.36	MktRet (t-5:t-8)
MktRet (t-9:t-12)	0.12	1.20			1.82	2.08	1.94	2.34	0.91	1.27			1.52	1.94	1.38	1.82	MktRet (t-9:t-12)
Flow (t-1)	0.41	9.25	0.45	12.14	0.39	11.18	0.42	13.36	0.39	8.71	0.41	10.88	0.38	9.91	0.40	11.40	Flow (t-1)
Flow (t-2)	0.01	0.24			0.03	0.63			0.00	0.04			0.02	0.51			Flow (t-2)
Flow (t-3)	0.12	2.85	0.11	3.35	0.10	3.31	0.10	3.30	0.09	2.66	0.08	2.42	0.09	3.08	0.10	3.38	Flow (t-3)
Flow (t-4)	0.01	0.29			0.10	2.38	0.07	2.14	0.14	3.61	0.16	4.84	0.14	3.63	0.13	3.74	Flow (t-4)
Flow (t-5:t-8)	0.04	2.63	0.04	3.18	0.01	0.70			0.00	0.06			0.00	-0.10			Flow (t-5:t-8)
Flow (t-9:t-12)	0.03	2.37	0.02	1.93	0.03	2.96	0.05	5.16	0.04	4.10	0.04	5.67	0.04	3.95	0.04	6.01	Flow (t-9:t-12)
Dec_h1_dummy	0.05	2.68	0.03	1.98	0.20	1.68			-0.10	-0.65			0.05	0.35			Dec_h1_dummy
Dec_h2_dummy	0.02	1.22			-1.88	-8.34	-2.01	-9.87	-1.91	-9.56	-1.87	-11.89	-2.03	-9.32	-2.06	-11.23	Dec_h2_dummy
Jan_h1_dummy	0.02	1.78			0.52	2.55	0.35	2.34	0.23	1.12	0.30	2.20	0.40	1.88	0.34	2.40	Jan_h1_dummy
Jan_h2_dummy	0.02	1.46			0.40	1.65			0.22	1.11	0.29	2.16	0.36	1.57	0.28	1.61	Jan_h2_dummy
Feb_dummy	0.04	3.08	0.03	2.66	0.49	3.36	0.26	4.16	0.12	0.80	0.16	2.33	0.28	1.88	0.25	4.17	Feb_dummy
Apr_dummy	0.00	0.29			-0.22	-1.40	-0.39	-3.54	-0.28	-1.70	-0.24	-2.35	-0.27	-1.57	-0.31	-2.88	Apr_dummy
May_dummy	0.02	1.51			0.17	1.28			-0.01	-0.07			0.07	0.48			May_dummy
Jun_h1_dummy	0.01	0.44			-0.08	-0.42	-0.24	-1.42	-0.16	-0.78			-0.14	-0.69			Jun_h1_dummy
Jun_h2_dummy	0.02	1.53			0.75	4.79	0.62	5.07	0.50	3.41	0.56	6.24	0.67	4.35	0.65	5.73	Jun_h2_dummy
Jul_h1_dummy	0.04	2.18			0.16	0.95			-0.14	-0.87			0.02	0.10			Jul_h1_dummy
Jul_h2_dummy	0.01	1.13			0.09	0.57			-0.07	-0.48			-0.02	-0.15			Jul_h2_dummy
Aug_dummy	0.02	1.63			0.20	1.62			0.03	0.19			0.12	0.94			Aug_dummy
Sep_dummy	-0.01	-0.82	-0.03	-2.03	-0.06	-0.43	-0.24	-2.55	-0.03	-0.16			-0.05	-0.34			Sep_dummy
Oct_dummy	0.01	0.43			0.10	0.73			0.00	0.03			0.05	0.41			Oct_dummy
Nov_dummy	0.03	2.08	0.02	1.14	0.31	2.16			0.02	0.14			0.17	1.18			Nov_dummy
Volatility(t)	-0.15	-0.29			29.21	7.18	28.07	6.84	27.21	6.87	26.50	6.71	30.22	7.65	30.25	7.61	Volatility(t)
Volatility(t-1)	-0.41	-1.07			-20.11	-5.13	-21.58	-6.37	-15.55	-4.08	-16.49	-4.71	-18.90	-4.94	-19.67	-5.42	Volatility(t-1)
Volatility(t-2)	0.64	1.61	0.60	1.64	-2.21	-0.70			-5.07	-1.78	-5.67	-2.13	-4.39	-1.49			Volatility(t-2)
Volatility(t-3)	0.06	0.17			-3.69	-0.56			-4.38	-0.65			-4.33	-0.62	-4.93	-0.71	Volatility(t-3)
Volatility(t-4)	-0.84	-2.35	-0.97	-2.98	-3.23	-0.97			0.67	0.23			-1.78	-0.57			Volatility(t-4)

TABLE 3: BROKER TYPE REGRESSIONS OF AGGREGATE NET FLOWS, BUY FLOWS, SELL FLOWS & TOTAL FLOWS

This table analyses normalised weekly net flows, buy flows and sell flows, allowing for different behaviour for internet investors and traditional full-service broker investors. Unnecessary regressors are eliminated using an iterative process that maximises the AIC information criterion. Variable definitions are the same as in Table II. For each flow type, internet results are on the left and traditional on the right. T-stats are calculated using the Newey-West method with four lags. The period of the regression is December 1996 to December 2002.

Variable	Net Flows				Buy Flows				Sell Flows				Total Flows				Variable
	Internet		Traditional		Internet		Traditional		Internet		Traditional		Internet		Traditional		
	Coef.	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	
p-value	0.000		0.000		0.000		0.000		0.000		0.000		0.000		0.000		p-value
adjR2	15.7%		39.8%		43.3%		57.8%		38.5%		49.6%		50.9%		55.9%		adjR2
AIC																	AIC
Constant	0.01	0.80	0.00	-0.15	-0.04	-0.45	-0.05	-0.48	0.07	0.58	-0.06	-0.53	0.03	0.31	-0.08	-0.75	Constant
MktRet (t)	-2.22	-4.34	-1.91	-6.45	-6.06	-2.37	-8.47	-4.53	7.28	2.68	4.94	2.04					MktRet (t)
MktRet (t-1)																	MktRet (t-1)
MktRet (t-2)							4.82	2.07	4.04	1.61	5.02	2.08	3.41	1.72	4.35	1.85	MktRet (t-2)
MktRet (t-3)			0.65	3.32													MktRet (t-3)
MktRet (t-4)																	MktRet (t-4)
MktRet (t-5:t-8)					3.95	3.27			2.94	2.30			1.92	1.62			MktRet (t-5:t-8)
MktRet (t-9:t-12)																	MktRet (t-9:t-12)
Flow (t-1)	0.16	2.33	0.43	10.36	0.33	6.69	0.43	9.83	0.35	7.07	0.37	5.66	0.37	7.72	0.42	7.68	Flow (t-1)
Flow (t-2)									0.10	2.13			0.12	1.87			Flow (t-2)
Flow (t-3)			0.13	2.48			0.11	2.36					0.09	1.65	0.11	2.47	Flow (t-3)
Flow (t-4)	-0.14	-1.32					0.12	2.90			0.10	2.48			0.10	2.64	Flow (t-4)
Flow (t-5:t-8)	0.07	2.18			0.03	2.05											Flow (t-5:t-8)
Flow (t-9:t-12)																	Flow (t-9:t-12)
Dec_h1_dummy									-0.32	-1.46							Dec_h1_dummy
Dec_h2_dummy			-0.03	-2.68	-1.75	-7.60	-2.04	-8.96	-1.71	-6.02	-2.24	-9.28	-1.81	-6.55	-2.24	-9.28	Dec_h2_dummy
Jan_h1_dummy							0.77	3.16					0.67	2.20	0.62	2.33	Jan_h1_dummy
Jan_h2_dummy			0.02	1.35	0.60	2.02	0.68	3.08	0.37	1.12			0.67	1.88	0.46	2.25	Jan_h2_dummy
Feb_dummy	0.04	2.76			0.78	4.86	0.50	5.08			0.32	2.36	0.41	3.21	0.42	4.00	Feb_dummy
Apr_dummy	-0.06	-2.25	-0.02	-1.64													Apr_dummy
May_dummy							0.25	2.52									May_dummy
Jun_h1_dummy			-0.03	-1.92													Jun_h1_dummy
Jun_h2_dummy					0.63	4.11	0.88	6.01	0.62	4.15	0.85	7.60	0.71	5.96	0.89	8.21	Jun_h2_dummy
Jul_h1_dummy	-0.05	-1.59															Jul_h1_dummy
Jul_h2_dummy					0.35	1.61											Jul_h2_dummy
Aug_dummy			0.02	1.53			0.20	2.04									Aug_dummy
Sep_dummy			-0.05	-2.97													Sep_dummy
Oct_dummy	-0.06	-2.35															Oct_dummy
Nov_dummy					0.38	2.85	0.39	3.19					0.27	2.79	0.25	2.27	Nov_dummy
Volatility(t)					18.01	4.45	20.36	3.25	17.16	3.74	26.67	4.86	21.16	4.90	24.83	4.11	Volatility(t)
Volatility(t-1)			0.67	1.84	-21.49	-4.12	-18.69	-4.72	-16.88	-3.39	-22.85	-6.28	-18.90	-3.84	-23.25	-6.43	Volatility(t-1)
Volatility(t-2)	1.94	2.93							-14.80	-3.02			-9.44	-1.60			Volatility(t-2)
Volatility(t-3)																	Volatility(t-3)
Volatility(t-4)			-0.80	-2.94			-6.11	-2.10	9.65	2.54							Volatility(t-4)

TABLE 4: INDIVIDUAL STOCK CROSS SECTIONAL ANALYSIS (ALL BROKERS): NET, BUY, SELL & TOTAL FLOWS

This table analyses weekly net flows, buy, sell and total flows. Individual weekly flows are calculated by summing buys and sells for each individual stock across all brokers in the sample each week. Net flows are defined as (buyflow – sellflow)/(buyflow + sellflow). Other flows are divided by weekly market capitalisation and are standardised cross sectionally so may be interpreted as z-scores. Buy, sell and total flows are in log terms. The period of the regression is September 1991 to December 2002. Contemporaneous and one period lagged stock returns are split into two separate variables, for positive and negative weeks eg StockRetPos = 0 if return is negative during that week and equals the return if the return is positive.

Variable	Net Flows		Buy Flows		Sell Flows		Total Flows		Variable
	All Brokers		All Brokers		All Brokers		All Brokers		
	Coef.	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	
Constant	0.00	0.35	-0.09	-18.05	-0.11	-20.23	-0.10	-21.25	Constant
PosStockRet (t)	**-1.39**	**-17.70**	1.43	12.61	**4.03**	**29.32**	3.01	28.20	**PosStockRet (t)**
NegStockRet (t)	**-1.86**	**-18.30**	**-3.23**	**-17.63**	-0.14	-0.99	-2.14	-14.33	**NegStockRet (t)**
PosStockRet (t-1)	-0.63	-9.91	0.16	1.75	1.36	12.51	0.65	7.56	**PosStockRet (t-1)**
NegStockRet (t-1)	-1.11	-12.40	-1.09	-8.76	0.58	4.56	-0.36	-3.29	**NegStockRet (t-1)**
StockRet (t-2)	-0.39	-9.72	-0.30	-4.76	0.27	4.24	-0.08	-1.31	StockRet (t-2)
StockRet (t-3)	-0.25	-7.58	-0.30	-7.73	0.03	0.47	-0.18	-4.42	StockRet (t-3)
StockRet (t-4)	-0.14	-4.72	-0.23	-5.30	-0.11	-2.29	-0.22	-5.17	StockRet (t-4)
StockRet (t-5:t-8)	-0.11	-5.85	-0.21	-7.75	-0.18	-5.06	-0.24	-8.11	StockRet (t-5:t-8)
StockRet (t-9:t-12)	0.00	-0.26	-0.09	-3.64	-0.27	-8.28	-0.19	-6.47	StockRet (t-9:t-12)
Flow (t-1)	0.26	33.73	0.32	29.82	0.26	26.64	0.31	25.68	Flow (t-1)
Flow (t-2)	0.10	33.15	0.14	36.22	0.12	34.21	0.13	39.94	Flow (t-2)
Flow (t-3)	0.08	22.65	0.09	26.34	0.10	28.88	0.10	27.53	Flow (t-3)
Flow (t-4)	0.06	16.48	0.07	20.85	0.08	19.17	0.08	20.24	Flow (t-4)
Flow (t-5:t-8)	0.03	21.08	0.04	20.72	0.05	25.72	0.04	20.94	Flow (t-5:t-8)
Flow (t-9:t-12)	0.03	22.98	0.03	23.00	0.04	25.89	0.03	20.72	Flow (t-9:t-12)
Stock_Volatility(t)	0.71	6.16	6.38	24.70	5.84	25.51	6.55	28.92	Stock_Volatility(t)
Stock_Volatility(t-1)	-0.16	-1.32	-2.15	-11.40	-1.65	-8.06	-2.23	-11.34	Stock_Volatility(t-1)
Stock_Volatility(t-2)	-0.21	-3.26	-1.58	-10.98	-1.06	-7.63	-1.56	-11.31	Stock_Volatility(t-2)
Stock_Volatility(t-3)	-0.42	-3.84	-1.18	-7.43	-0.51	-3.07	-0.99	-5.96	Stock_Volatility(t-3)
Stock_Volatility(t-4)	-0.47	-4.93	-1.58	-10.76	-0.94	-5.95	-1.51	-10.76	Stock_Volatility(t-4)

TABLE 5: INDIVIDUAL STOCK CROSS SECTIONAL ANALYSIS (INTERNET VS TRADITIONAL): NET, BUY, SELL & TOTAL FLOWS

This table analyses weekly net flows, buy flows and sell flows separately for internet and traditional full-service brokerage clients. Individual weekly flows are calculated by summing buys and sells for each individual stock across all brokers in the sample each week. Net flows are defined using raw flows as (buyflow – sellflow)/(buyflow + sellflow). Buy, sell and total flows are divided by weekly market capitalisation and are standardised cross sectionally so may be interpreted as z-scores. The period of the regression is December 1996 to December 2002.

Variable	Net Flows				Buy Flows				Sell Flows				Total Flows			
	Internet		Traditional		Internet		Traditional		Internet		Traditional		Internet		Traditional	
	Coef.	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*	*Coef.*	*T-stat*
Constant	0.01	1.66	0.00	-0.02	-0.11	-17.29	-0.10	-15.30	-0.12	-14.41	-0.11	-14.72	-0.11	-16.92	-0.10	-16.23
PosStockRet (t)	**-2.61**	**-18.61**	**-0.93**	**-9.18**	0.73	4.63	1.89	10.84	**4.79**	**23.50**	**3.80**	**21.24**	2.95	18.63	3.12	19.98
NegStockRet (t)	**-3.27**	**-14.84**	**-1.37**	**-9.10**	**-4.69**	**-19.12**	**-3.21**	**-21.00**	0.49	2.23	-0.81	-3.52	-3.01	-17.17	-2.42	-15.61
PosStockRet (t-1)	-0.87	-5.57	-0.72	-9.56	-0.02	-0.14	-0.06	-0.68	1.39	8.69	1.38	9.83	0.37	3.65	0.57	6.88
NegStockRet (t-1)	-1.55	-7.39	-0.98	-10.21	-1.57	-8.10	-1.17	-8.00	0.24	1.51	0.28	2.10	-0.89	-6.51	-0.50	-3.80
StockRet (t-2)	-0.44	-6.14	-0.41	-9.25	-0.42	-5.30	-0.48	-6.94	-0.05	-0.76	0.18	2.57	-0.40	-6.30	-0.22	-3.53
StockRet (t-3)	-0.13	-2.25	-0.27	-5.83	-0.20	-3.25	-0.41	-7.10	-0.31	-3.80	-0.03	-0.63	-0.27	-4.31	-0.28	-5.99
StockRet (t-4)	-0.01	-0.18	-0.15	-4.38	-0.04	-0.82	-0.22	-3.57	-0.35	-5.62	-0.07	-1.26	-0.14	-2.89	-0.17	-2.63
StockRet (t-5:t-8)	-0.04	-1.09	-0.09	-4.40	-0.06	-2.22	-0.18	-5.12	-0.31	-6.75	-0.15	-3.53	-0.16	-5.44	-0.19	-4.87
StockRet (t-9:t-12)	0.08	2.33	0.01	0.43	0.04	1.91	-0.05	-1.75	-0.28	-5.78	-0.21	-4.87	-0.07	-2.82	-0.13	-3.84
Flow (t-1)	0.23	21.25	0.26	49.95	0.34	23.39	0.33	53.27	0.24	19.92	0.28	32.71	0.33	21.75	0.34	39.57
Flow (t-2)	0.08	16.86	0.11	33.83	0.13	25.77	0.15	30.40	0.12	21.54	0.12	22.46	0.14	34.03	0.14	34.54
Flow (t-3)	0.08	13.90	0.08	18.08	0.10	18.86	0.09	20.99	0.10	20.31	0.10	20.98	0.11	20.89	0.10	23.72
Flow (t-4)	0.05	9.32	0.06	11.39	0.07	11.70	0.07	14.48	0.08	20.19	0.08	14.57	0.07	16.00	0.08	17.73
Flow (t-5:t-8)	0.03	18.62	0.03	16.44	0.04	18.60	0.04	15.56	0.05	23.02	0.04	17.87	0.04	17.30	0.04	19.10
Flow (t-9:t-12)	0.02	13.00	0.03	17.49	0.03	13.18	0.03	16.46	0.04	17.12	0.04	21.16	0.03	12.90	0.03	17.63
Stock_Volatility(t)	1.28	5.62	0.42	3.18	6.52	17.25	6.20	19.18	5.38	12.28	5.74	15.98	6.27	15.52	6.38	18.22
Stock_Volatility(t-1)	-0.14	-0.66	0.03	0.20	-2.18	-9.68	-2.00	-10.55	-1.55	-7.54	-1.85	-8.45	-2.30	-9.58	-2.40	-11.92
Stock_Volatility(t-2)	-0.31	-2.78	-0.20	-1.82	-1.70	-11.09	-1.65	-7.28	-0.79	-3.94	-1.08	-5.01	-1.53	-9.08	-1.49	-6.81
Stock_Volatility(t-3)	-0.36	-2.51	-0.45	-4.18	-1.42	-8.22	-1.32	-7.74	-0.55	-2.74	-0.93	-4.42	-1.40	-8.33	-1.35	-8.03
Stock_Volatility(t-4)	-0.58	-3.83	-0.32	-2.54	-1.65	-9.96	-1.56	-6.36	-0.84	-3.96	-1.21	-5.15	-1.43	-9.22	-1.47	-7.28

TABLE 6: BROKER-PAIR CORRELATION ANALYSIS

This table shows the average pairwise correlation between brokers for four different broker types. Brokers are divided into 8 Internet brokers, 39 full-service retail brokers, 12 mixed brokers & 11 institutional brokers. The first row of the table is the average pairwise correlation and other statistics for the 8x7/2 = 28 internet broker pairs, the fifth row shows the statistics for the 39x8 = 312 full-service / internet broker pairs. The second column shows the correlation across the pairs, the third column the percentage of pairwise correlations that are greater than zero. The percentage of t-statistics that are greater than or less than 2 are given in the next two columns, and the number of broker pairs used in the sample are given in the final column. The sample period for the analysis is September 1991 to December 2002. To calculate the correlations for a single broker pair, the correlation between net flows for each week is taken for the cross-section of all index stocks, this is repeated for each week, then the time series of the correlation is taken to get a point estimate and calculate t-statistics.

Broker Type 1	Broker Type 2	Average Pairwise Correlation	Stdev Pairwise Correlation	Percentage of Positive Correlation Pairs	Percentage of pairs with t-stats > 2	Percentage of pairs with t-stats <- 2	Number of BrokerPairs
Internet	Internet	0.229	0.045	100%	100%	0%	28
Full-Service	Full-Service	0.159	0.040	100%	100%	0%	741
Mixed	Mixed	0.021	0.033	67%	32%	68%	66
Insto	Insto	-0.003	0.013	36%	13%	87%	55
Internet	Full-Service	0.156	0.042	100%	100%	0%	312
Internet	Mixed	0.059	0.038	93%	85%	15%	96
Internet	Insto	-0.023	0.044	27%	23%	77%	88
Full-Service	Mixed	0.060	0.043	90%	77%	23%	468
Full-Service	Insto	-0.008	0.032	31%	24%	76%	429
Mixed	Insto	-0.004	0.031	36%	15%	85%	132

TABLE 7: BROKER-PAIR CORRELATION ANALYSIS: TOP 50 STOCKS ONLY

This table shows the average pairwise correlation between brokers for four different broker types for the Top 50 stocks only. Brokers are divided into 8 Internet brokers, 39 full-service retail brokers, 12 mixed brokers & 11 institutional brokers. The first row of the table is the average pairwise correlation and other statistics for the 8x7/2 = 28 internet broker pairs, the fifth row shows the statistics for the 39x8 = 312 full-service / internet broker pairs. The second column shows the correlation across the pairs, the third column the percentage of pairwise correlations that are greater than zero. The percentage of t-statistics that are greater than or less than 2 are given in the next two columns, and the number of broker pairs used in the sample are given in the final column. The sample period for the analysis is September 1991 to December 2002. To calculate the correlations for a single broker pair, the correlation between net flows for each week is taken for the cross-section of all index stocks, this is repeated for each week, then the time series of the correlation is taken to get a point estimate and calculate t-statistics.

Broker Type 1	Broker Type 2	Average Pairwise Correlation	Stdev Pairwise Correlation	Percentage of Positive Correlation Pairs	Percentage of pairs with t-stats > 2	Percentage of pairs with t-stats <- 2	Number of BrokerPairs
Internet	Internet	0.439	0.073	100%	100%	0%	28
Full-Service	Full-Service	0.235	0.056	100%	100%	0%	741
Mixed	Mixed	0.032	0.046	68%	38%	62%	66
Insto	Insto	-0.011	0.016	24%	4%	96%	55
Internet	Full-Service	0.270	0.062	100%	100%	0%	312
Internet	Mixed	0.125	0.080	100%	92%	8%	96
Internet	Insto	-0.024	0.058	30%	19%	81%	88
Full-Service	Mixed	0.091	0.066	98%	82%	18%	468
Full-Service	Insto	-0.011	0.041	25%	19%	81%	429
Mixed	Insto	-0.011	0.030	35%	11%	89%	132

TABLE 8 – VAR FOR MARKET RETURNS AND FLOWS

This table reports parameter estimates for a VAR containing weekly aggregate market returns (in %) and aggregate weekly raw flows. The process is estimated for all brokers, and separately for internet and full-service brokers. The parameters in the table correspond to the reduced VAR system described in section V part A. The bottom section of the table contains the results of the Granger causality test with the associated F-statistics & p-values.

Variable	All Investors (1991-2002)				Internet Investors (1996-2002)				Full-Service Investors (1996-2002)			
	Flow (t)	*tstat*	*Return (t)*	*tstat*	*Flow (t)*	*tstat*	*Return (t)*	*tstat*	*Flow (t)*	*tstat*	*Return (t)*	*tstat*
Constant	-0.002	-0.18	-0.0031	-1.12	-0.001	-0.02	-0.0028	-0.73	-0.033	-1.74	-0.0021	-0.55
Flow (t-1)	0.451	8.58	**0.0215**	**2.66**	0.221	3.09	**0.0096**	**1.42**	0.380	6.58	**0.0329**	**2.86**
Flow (t-2)	0.129	2.26	**-0.0235**	**-2.23**	-0.045	-0.62	**-0.0054**	**-0.64**	0.156	2.09	**-0.0489**	**-3.13**
Flow (t-3)	0.155	3.11	0.0097	1.17	-0.019	-0.26	0.0073	0.92	0.078	1.09	0.0243	2.01
StockRet (t-1)	-0.584	-1.88	0.0412	0.89	-0.290	-0.65	-0.0228	-0.36	-0.298	-0.81	0.0574	0.80
StockRet (t-2)	-0.023	-0.08	0.0001	0.00	-0.378	-0.57	-0.0068	-0.12	0.351	0.93	-0.0754	-1.36
StockRet (t-3)	0.254	1.04	0.0549	1.19	-0.283	-0.47	0.0305	0.50	0.667	2.49	0.0160	0.26

Variable	All Investors				Internet Investors				Traditional Investors			
	Flow (t)	*p value*	*Return (t)*	*p value*	*Flow (t)*	*tstat*	*Return (t)*	*p value*	*Flow (t)*	*p value*	*Return (t)*	*p value*
Granger F-Stat Flow (t-1)	117.78	0.000	6.26	0.023	5.67	0.001	0.92	0.430	31.97	0.000	5.01	0.002
Granger F-Stat Return (t-1)	1.907	0.127	0.809	0.489	0.712	0.545	0.111	0.954	2.386	0.069	1.046	0.373

TABLE 9 – VAR ESTIMATES FOR CROSS-SECTIONAL RETURNS AND FLOWS

This table reports parameter estimates for the cross sectional flow VAR. The process is estimated for all investors and separately for internet and traditional full-service investors. The parameters in the table correspond to the reduced form VAR system with 12 lags. For each parameter I report the estimated coefficient value and the t-stat of the coefficient estimated using Newey West standard errors. The VAR is estimated from September 1991 to Dec 2002.

Variable	Full-Service + Internet: 1991-2002				Internet: 1996-2002				Full-Service: 1996-2002			
	Flow (t)	*tstat*	*Return (t)*	*tstat*	*Flow (t)*	*tstat*	*Return (t)*	*tstat*	*Flow (t)*	*tstat*	*Return (t)*	*tstat*
Constant	0.001	0.47	0.001	0.52	0.03	4.67	0.008	1.72	0.00	-0.42	0.000	0.07
StockRet (t-1)	-0.035	-15.72	-0.076	-10.36	-0.05	-7.39	-0.052	-7.31	-0.04	-14.59	-0.049	-6.92
StockRet (t-2)	-0.017	-10.51	-0.014	-2.84	-0.02	-6.40	-0.003	-0.46	-0.02	-9.85	0.000	0.00
StockRet (t-3)	-0.013	-8.60	0.013	2.57	-0.01	-3.03	0.017	2.21	-0.02	-7.01	0.024	3.29
StockRet (t-4)	-0.009	-6.43	0.019	3.71	-0.01	-2.51	0.020	2.48	-0.01	-5.69	0.030	4.50
StockRet (t-5)	-0.008	-5.86	0.010	2.08	-0.01	-1.89	0.011	1.37	-0.01	-7.25	0.017	2.95
StockRet (t-6)	-0.008	-5.04	0.010	2.22	0.00	-0.92	0.008	1.11	-0.01	-3.59	0.010	1.83
StockRet (t-7)	-0.005	-3.76	0.013	2.85	0.00	-1.15	0.007	0.99	-0.01	-2.91	0.013	2.15
StockRet (t-8)	-0.004	-3.04	0.015	3.31	0.00	-0.07	0.016	2.32	0.00	-1.70	0.018	2.94
StockRet (t-9)	-0.004	-3.06	0.014	2.71	0.00	-0.22	0.014	1.90	0.00	-3.70	0.013	1.85
StockRet (t-10)	0.000	-0.25	0.013	2.48	0.00	1.53	0.022	3.02	0.00	-1.06	0.022	3.69
StockRet (t-11)	0.001	0.45	0.012	2.81	0.00	-0.10	0.013	1.76	0.00	1.07	0.014	2.15
StockRet (t-12)	0.000	0.27	0.009	2.17	0.01	1.86	0.014	2.13	0.00	1.71	0.014	2.48
Flow (t-1)	0.262	34.18	**0.036**	**4.29**	0.25	23.00	**-0.010**	**-1.06**	0.26	46.26	**0.065**	**7.25**
Flow (t-2)	0.101	28.81	**0.026**	**3.59**	0.08	15.14	**0.018**	**1.99**	0.11	29.36	**0.034**	**3.28**
Flow (t-3)	0.074	20.06	0.013	1.85	0.08	11.53	0.015	1.47	0.08	16.89	0.017	1.62
Flow (t-4)	0.056	15.04	0.006	0.86	0.05	6.85	0.006	0.65	0.06	10.31	0.003	0.35
Flow (t-5)	0.043	12.10	0.000	-0.01	0.04	5.94	-0.010	-1.01	0.04	8.48	-0.008	-0.93
Flow (t-6)	0.033	8.09	-0.006	-0.81	0.03	4.97	0.001	0.12	0.04	7.98	-0.007	-0.68
Flow (t-7)	0.023	7.22	0.003	0.39	0.02	3.44	-0.004	-0.42	0.02	4.34	0.009	0.90
Flow (t-8)	0.029	8.09	-0.012	-1.60	0.03	4.38	-0.007	-0.68	0.02	4.89	-0.024	-2.05
Flow (t-9)	0.032	9.04	0.000	0.02	0.03	6.99	0.004	0.44	0.03	6.97	-0.001	-0.12
Flow (t-10)	0.030	7.89	0.005	0.67	0.03	5.79	-0.009	-0.93	0.03	6.19	0.012	1.39
Flow (t-11)	0.024	7.28	-0.017	-2.33	0.02	4.86	-0.022	-2.49	0.02	4.74	-0.013	-1.29
Flow (t-12)	0.028	8.23	-0.009	-1.24	0.02	5.37	0.004	0.58	0.03	8.27	-0.021	-2.39

TABLE 10 – VAR ESTIMATES FOR CROSS-SECTIONAL RETURNS AND FLOWS: LARGE AND SMALL STOCKS

This table reports parameter estimates for the cross sectional flow VAR estimated separately for large and small stocks. The process is estimated for all investors and separately for internet and full-service investors. The parameters in the table correspond to the reduced form VAR system with 12 lags. For each parameter I report the estimated coefficient value and the t-stat of the coefficient estimated using Newey West standard errors. The VAR is estimated from September 1991 to Dec 2002. Only the return equations are reported, with the flow equations omitted to save space.

Variable	**Full-Service + Internet: 1991-2002**				**Internet: 1996-2002**				**Full-Service: 1996-2002**			
	Large (Top 100)		Small (Ex 100)		Large (Top 100)		Small (Ex 100)		Large (Top 100)		Small (Ex 100)	
	Return (t)	tstat	Return (t)	tstat	Return (t)	tstat	Return (t)	tstat	Return (t)	tstat	Return (t)	tstat
Constant	-0.026	-5.28	0.011	3.75	-0.026	-2.74	0.015	2.39	-0.021	-3.15	0.011	3.64
StockRet (t-1)	-0.103	-13.28	-0.072	-8.23	-0.100	-9.20	-0.043	-4.02	-0.098	-12.02	-0.045	-5.66
StockRet (t-2)	-0.017	-2.53	-0.019	-3.04	-0.014	-1.41	-0.006	-0.64	-0.003	-0.34	-0.003	-0.45
StockRet (t-3)	0.008	1.11	0.008	1.29	0.007	0.76	0.020	1.85	0.014	1.54	0.021	2.28
StockRet (t-4)	0.013	1.72	0.019	3.20	0.011	1.07	0.028	3.27	0.016	1.52	0.032	4.65
StockRet (t-5)	0.012	1.49	0.007	1.31	0.014	1.35	0.001	0.09	0.023	2.38	0.011	1.75
StockRet (t-6)	0.008	1.22	0.010	1.73	-0.001	-0.11	0.005	0.61	0.005	0.56	0.012	1.82
StockRet (t-7)	0.004	0.50	0.016	2.81	-0.007	-0.51	0.009	0.94	0.005	0.47	0.015	1.95
StockRet (t-8)	0.011	1.51	0.020	3.54	0.005	0.43	0.025	2.76	0.013	1.24	0.024	3.40
StockRet (t-9)	0.018	2.43	0.017	2.83	0.025	2.70	0.007	0.75	0.011	1.24	0.015	1.96
StockRet (t-10)	0.018	3.07	0.007	1.14	0.029	3.22	0.006	0.62	0.018	1.90	0.018	2.44
StockRet (t-11)	0.008	1.26	0.012	2.32	0.014	1.28	0.009	1.11	0.010	1.05	0.010	1.47
StockRet (t-12)	0.011	2.09	0.008	1.49	0.006	0.62	0.016	1.94	0.015	2.12	0.015	2.16
Flow (t-1)	0.077	4.79	0.019	1.78	0.001	0.05	-0.028	-2.52	0.116	5.50	0.052	4.56
Flow (t-2)	0.001	0.08	0.022	2.60	0.018	1.03	0.013	1.28	0.006	0.23	0.035	3.21
Flow (t-3)	0.048	3.55	0.007	0.76	0.023	1.07	0.006	0.44	0.065	3.60	0.004	0.34
Flow (t-4)	-0.015	-1.05	0.012	1.34	0.011	0.66	0.009	0.77	-0.026	-1.58	0.013	1.36
Flow (t-5)	0.021	1.16	-0.007	-0.65	-0.013	-0.60	0.001	0.07	0.011	0.49	-0.010	-0.92
Flow (t-6)	-0.022	-1.57	-0.001	-0.08	0.005	0.26	-0.009	-0.66	-0.030	-1.46	0.006	0.49
Flow (t-7)	0.033	1.97	-0.003	-0.33	-0.002	-0.07	0.012	0.95	0.052	2.38	-0.002	-0.17
Flow (t-8)	-0.020	-1.32	-0.006	-0.64	0.006	0.28	-0.001	-0.08	-0.055	-2.97	-0.014	-1.09
Flow (t-9)	0.018	1.07	-0.006	-0.58	0.052	2.50	-0.018	-1.20	0.005	0.21	-0.007	-0.55
Flow (t-10)	-0.008	-0.51	0.006	0.57	-0.010	-0.44	-0.019	-1.46	-0.013	-0.70	0.016	1.38
Flow (t-11)	-0.031	-2.24	-0.016	-1.68	-0.012	-0.55	0.001	0.07	-0.012	-0.58	-0.014	-1.04
Flow (t-12)	-0.007	-0.47	-0.002	-0.21	0.002	0.10	0.008	0.62	-0.018	-0.88	-0.024	-2.26

TABLE 11 – PORTFOLIO RETURNS BASED ON NET FLOWS BY BROKERAGE TYPE

This table reports the portfolio returns for long and short portfolios constructed based on the net flows of each broker type. The zero cut-off portfolios assign a stock to the long portfolio is the brokerage type is a net buyer each week, the 0.25 cut-off assigns a stock to the long portfolio only if the net flow ratio is >0.25. Results are reported in basis points per week for the various portfolios.

Panel A: Zero Cut-off

Broker	Portfolio	t-5	t-4	t-3	t-2	t-1	t	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10
Big	Long	34.0	34.2	31.5	33.2	45.5	60.7	10.4	13.4	12.8	13.6	14.7	13.3	13.8	14.1	13.8	12.7
(1991-2002)	Short	0.4	-1.3	1.7	-1.3	-12.6	-28.2	15.8	12.3	12.9	10.9	9.5	11.2	12.4	11.6	11.7	13.2
	Difference	**33.7**	**35.5**	**29.8**	**34.4**	**58.1**	**88.9**	**-5.4**	**1.1**	**-0.1**	**2.7**	**5.2**	**2.2**	**1.4**	**2.5**	**2.1**	**-0.5**
	t-stat	10.6	10.9	8.7	9.7	15.2	21.9	-1.8	0.4	0.0	0.9	1.7	0.7	0.5	0.9	0.7	-0.2
Mix	Long	23.5	19.5	22.7	21.4	22.4	27.2	16.7	14.3	16.9	12.4	12.3	16.0	15.5	17.0	13.4	13.5
(1991-2002)	Short	8.4	12.2	9.0	9.3	7.6	2.7	9.8	10.5	8.2	11.3	10.9	8.8	9.8	8.0	12.1	11.4
	Difference	**15.0**	**7.3**	**13.7**	**12.1**	**14.8**	**24.5**	**6.9**	**3.9**	**8.7**	**1.1**	**1.3**	**7.1**	**5.7**	**9.0**	**1.4**	**2.2**
	t-stat	5.2	2.6	4.9	4.2	5.1	8.2	2.4	1.4	3.1	0.4	0.5	2.5	2.1	3.2	0.5	0.8
Full-Service	Long	-4.6	-3.6	-8.1	-10.3	-23.4	-25.0	23.7	18.6	13.2	14.6	14.0	11.5	13.7	10.1	11.6	12.7
(1991-2002)	Short	36.0	33.9	37.8	38.7	50.4	51.2	2.5	7.0	12.0	10.1	10.2	13.0	11.8	16.0	13.0	11.7
	Difference	**-40.6**	**-37.5**	**-45.9**	**-49.1**	**-73.8**	**-76.2**	**21.2**	**11.6**	**1.2**	**4.4**	**3.8**	**-1.5**	**1.9**	**-6.0**	**-1.4**	**1.0**
	t-stat	-13.8	-12.9	-15.2	-17.1	-22.2	-21.2	7.0	3.9	0.4	1.6	1.3	-0.5	0.7	-2.1	-0.5	0.4
Int	Long	-13.6	-14.5	-17.7	-29.3	-45.7	-85.5	5.8	4.1	4.5	0.1	3.4	3.0	0.7	-0.2	-0.3	-1.3
(1996-2002)	Short	42.5	44.7	48.2	60.9	85.7	139.2	8.9	14.5	11.3	13.9	14.2	15.3	15.3	15.3	14.3	15.8
	Difference	**-55.8**	**-59.2**	**-65.7**	**-90.8**	**-132.1**	**-224.5**	**-3.0**	**-12.0**	**-6.9**	**-13.8**	**-11.7**	**-13.2**	**-14.5**	**-15.7**	**-14.6**	**-16.9**
	t-stat	-12.3	-12.7	-14.9	-19.2	-26.8	-33.8	-0.6	-2.6	-1.6	-3.1	-2.4	-2.9	-3.1	-3.4	-3.4	-3.9

Panel B: Cut-off of Net flow = 0.25

Broker	Portfolio	t-5	t-4	t-3	t-2	t-1	t	t+1	t+2	t+3	t+4	t+5	t+6	t+7	t+8	t+9	t+10
Big	Long	31.2	30.0	23.9	28.4	38.9	52.0	0.2	5.5	11.8	11.4	5.7	13.3	10.8	10.5	12.7	8.8
	Short	-16.7	-18.8	-10.3	-16.8	-29.6	-46.9	9.9	7.1	2.8	4.4	-0.7	8.2	5.2	5.3	5.8	10.8
	Difference	47.9	48.8	34.1	45.2	68.6	98.8	-9.7	-1.6	8.9	6.9	6.5	5.2	5.6	5.2	6.9	-2.0
	t-stat	8.6	8.1	5.8	7.3	10.9	14.8	-1.8	-0.3	1.8	1.3	1.1	1.0	1.1	1.0	1.3	-0.4
Mix	Long	23.9	18.7	18.6	14.3	16.2	17.5	15.3	11.7	19.4	12.5	14.4	16.9	13.6	16.1	16.3	13.6
	Short	3.4	2.9	4.2	3.2	-5.7	-14.7	4.6	3.2	7.1	5.7	4.7	4.7	7.1	3.0	9.9	6.6
	Difference	20.5	15.8	14.4	11.1	21.9	32.2	10.7	8.5	12.3	6.8	9.7	12.2	6.5	13.1	6.4	6.9
	t-stat	4.6	3.6	3.5	2.4	5.4	7.1	2.4	1.9	2.8	1.6	2.2	2.7	1.5	3.0	1.5	1.7
Nor	Long	-7.1	-7.2	-15.2	-20.0	-40.5	-46.7	30.2	24.7	18.7	19.9	18.7	12.6	17.4	11.7	15.4	13.5
	Short	45.3	43.2	46.1	48.7	58.7	58.3	-0.2	7.6	12.9	12.8	12.6	16.7	13.5	18.5	14.0	13.5
	Difference	-52.4	-50.4	-61.3	-68.7	-99.2	-105.0	**30.4**	**17.1**	5.8	7.1	6.0	-4.1	3.9	-6.8	1.4	0.0
	t-stat	-14.6	-13.4	-16.5	-19.0	-24.7	-24.2	**7.9**	**4.4**	1.5	1.9	1.6	-1.1	1.0	-1.9	0.4	0.0
Int	Long	-14.4	-21.7	-24.3	-36.9	-67.7	-132.4	9.2	9.2	7.3	2.0	6.3	3.9	4.3	1.9	3.0	-2.6
	Short	52.6	53.1	55.6	66.1	95.2	150.9	12.5	17.8	17.1	18.5	17.1	19.7	19.5	21.2	15.2	21.5
	Difference	-66.9	-74.8	-79.8	-103.8	-163.9	-283.5	-2.7	-11.4	-9.5	-16.2	-11.3	-16.1	-15.1	-18.8	-12.1	-23.5
	t-stat	-12.9	-14.6	-15.1	-19.7	-27.0	-33.5	-0.5	-2.1	-1.9	-3.1	-2.0	-3.0	-2.9	-3.5	-2.4	-4.6