Retail Investor Networks and Risk Taking

Henk Berkman

Marco Eugster*

The University of Auckland Business School

28 January 2015

Abstract

In contrast to a recent theory proposing that social interaction encourages retail investors to trade high-risk stocks, we show that social interaction has a moderating effect. Lower risk ("conservative") investors increase their purchases in high-volatility and high-skewness stocks following social interaction while high risk ("speculative") individuals decrease their purchases. However, this moderating effect appears to be significantly stronger for speculators. Our results are consistent with a number of studies from the experimental economics literature which suggest that group interaction leads to more risk-averse and rational behaviour.

*Corresponding author. Email: marco.eugster@auckland.ac.nz

This version is preliminary. Please do not cite without permission.

1 Introduction

Social interaction among retail investors is common. Individuals recount trading experiences, discuss strategies and share new ideas. Our understanding of the effect that this kind of social interaction has on subsequent investment behaviour is still very limited, however. One possibility is that interaction encourages risky trading strategies, exacerbates behavioural biases and harms portfolio performance (Han & Hirshleifer, 2012). An alternative view is that interaction acts as a moderating factor on investor behaviour, promoting more rational decision making and pushing investors away from the extremes (Kugler, Kausel, & Kocher, 2012).

In this paper, we examine these competing hypotheses in the context of risk-taking and find that social interaction has a moderating impact on investor behaviour. When we split investors into two groups based on their risk-taking behaviour during the pre-interaction period, we find that both conservative and speculative investors converge to the more average risk-taking displayed by their network peers. However, the decrease in risk-taking displayed by speculators following social interaction is significantly stronger than the increase in risk-taking displayed by conservatives. This suggests that speculative investors may be particularly prone to the moderating effect of social interaction.

Our empirical tests are based on a sample of Finnish retail investors between the years 1997 through 2011. In constructing our network of connected investors, we rely on a method recently introduced to the finance literature by Ozsoylev, Walden, Yavuz and Bildik (2014). Two investors are defined as being socially connected if, on regular occasions, they buy the same stock within a short period of time. This approach allows us to classify any pair of investors as being either socially connected or unconnected based solely on their actual stock market investment behaviour. We introduce a number of additional controls in order to improve the methodology's precision. These are important because they allow us to better differentiate between the true localised social interaction we intend to capture and reactions to centralised information diffusion or common trading strategies.

We first validate our proxy for social interaction by examining the exogenous characteristics of the pairs of investors we identify as being socially connected. We find that social connections are heavily clustered around individuals whom we would expect to have a high chance of being connected. For example, 26% of the social connections we observe occur between investors who reside less than 5 normalised-equivalent kilometres apart. Based on the investors in our sample, we would expect only 2% of connections to fall within this range. Connections are also clustered around investors with familial relationships, who are of similar age and who speak the same language. These tests also contribute to the existing literature by highlighting the channels of social interaction among investors. For example, they demonstrate the importance of familial ties in investment decision making.

Our main empirical tests proceed in the form of an event study. For a given investor, we define year zero as the first year during which we identify the investor as being socially connected to at least one other individual. We measure exposure to high-risk trading strategies by the proportion of an investor's annual trading activity that is allocated to stocks falling within the highest volatility, idiosyncratic volatility and skewness quintiles. We calculate a market-adjusted, or "abnormal", proportion of trading in these stocks (which we term APROP) by subtracting the average proportion calculated across other investors in the market with similar trading frequencies. We then track the development of APROP for a given investor in event time.

We also track the APROP calculated across the individuals comprising an investor's year-zero peer network. We present strong evidence that social interaction acts as a moderating factor on risk-taking by investors. An increase by one percentage point in the abnormal proportion that an individual's social network trades in high-volatility stocks during the pre-connection period is associated with an increase of 0.29 percentage points in the individual's abnormal portion traded in high-volatility stocks during the year of first connection. Closer inspection reveals that the moderating effect of social interaction has an asymmetric effect on speculative and conservative investors. The tendency of investors who display a positive value of APROP in the pre-connection period ("speculators") to reduce their purchases of high-risk stocks following social interaction appears to be greater than the tendency of investors who display a negative value of APROP in the pre-connection period ("conservatives") to increase their purchases of high-risk stocks. This result is consistent with the findings by Masclet, Colombier, Denant-Boemont and Loheac (2009) in an experimental setting.

A relatively recent literature documents the existence of peer effects in various investment contexts. Social interaction has been shown both to encourage stock market participation (Brown, Ivković, Smith, & Weisbenner, 2008; Guiso & Jappelli, 2005; Liu, Meng, You, & Zhao, 2013) and also to promote correlated securities choices among connected investors (Hvide & Östberg, 2013; Ivković & Weisbenner, 2007; Shive, 2010). Investors also appear to be responsive to outcomes experienced by their peers (Hellström, Zetterdahl, & Hanes, 2013; Kaustia & Knüpfer, 2012; Lu, 2012). The influence of social interaction has been documented in other investment-related settings including institutional investor behaviour (Hong, Kubik, & Stein, 2005; Pool, Stoffman, & Yonker, 2013), security analysts (Horton & Serafeim, 2009), insider trading (Ahern, 2014) and retirement savings choices (Duflo & Saez, 2002).

In contrast, evidence as to how social interaction might influence investors' risky trading behaviour is much scarcer and less consistent. A theoretical model by Han and Hirshleifer (2012) proposes that social interaction involves an inherent bias in the transmission process.[1] The bias is referred to as "self-enhancing transmission bias", reflecting the notion that individuals are more likely to recount positive investment outcomes than negative ones. Transmission receivers do not fully discount this bias and thus overweight the value of the strategy. The higher a strategy's return variance, the stronger the bias. As a result, strategies with high volatility and high skewness propagate among connected investors. Such strategies can have a detrimental impact on portfolio performance, especially if they lead to underdiversification or systematic purchases of overvalued securities. Few empirical studies have directly tested the predictions from the Han and Hirshleifer (2012) model. Those that do have tended to focus on specific investor groups. For example, consistent with the model's precitions, Simon and Heimer (2012) find that foreign exchange

[1] The model by Han and Hirshleifer (2012) follows an extensive literature on the role of social learning in economic outcomes. Early sequential decision models describe how informational cascades and herding may arise when individuals base their decisions, at least in part, on the actions of others (A. V. Banerjee, 1992; Bikhchandani, Hirshleifer, & Welch, 1992; Welch, 1992). Ellison and Fudenberg (1993; 1995) and Banerjee and Fudenberg (2004) focus on the nature of the transmission process and show that product decisions based on naïve rules of thumb and/or word-of-mouth communication can lead to fairly efficient long-run social outcomes. Conversely, Cao, Han and Hirshleifer (2011) suggest that information externalities limit efficient decision-making and make individuals prone to mistaken informational cascades.

traders are more likely to initiate interaction and react more strongly to interaction following positive performance. Moreover, trading intensity and return variance experienced by traders in the network rise over time. It is unclear whether conclusions drawn for foreign exchange traders, who are often perceived as a particularly active type of trader, are valid for retail stock investors, however.

A second, conflicting, strand of evidence comes from the experimental economics literature. The balance of evidence in this area points towards a role for groups as a moderator of behaviour. In a recent review on group decision making, Kugler et al., (2012) conclude that group decisions more closely approximate the predicted behaviour of a rational economic agent than individual decisions. In an experimental setting, Masclet et al., (2009) find that groups are more likely to choose safe lotteries than individuals. Moreover, less risk averse individuals are found to be more willing to shift their decision to conform to the group average. Other experiments suggest that social interaction and/or information sharing lead to closer adherence to the principles of Markowitz portfolio selection theory, (Rockenbach, Sadrieh, & Mathauschek, 2007), improves underdiversification and reduces risk, (Baghestanian, Gortner, & Van der Weele, Joel J, 2014) and reduces the likelihood of bubbles and crashes in an asset market (Cheung & Palan, 2012).

Given the significant role that social interaction has been shown to play in the investment process, understanding its influence on risk-taking is of major importance. The key contribution which our study makes to the literature is to empirically investigate the two competing predictions from the finance and experimental economics literatures discussed above. Our results suggest a role for social interaction that has received little or no consideration in the finance literature: as a moderator of more extreme investor behaviour. This notion potentially has important implications. To the extent that a tendency to overinvest in high-risk and high-skewness stocks hurts portfolio performance, close interaction among retail investors might actually be beneficial.

2 Data and empirical design

2.1 Data

This study uses daily data on stock trading by Finnish individual investors on the Nasdaq OMX Helsinki exchange from January 1997 through December 2011. To trade on the exchange, investors must register with Euroclear. Each investor obtains a single unique Euroclear account which aggregates the trades made by that investor, even if they are made through different brokers. We aggregate the trades which an investor makes in a given stock across each day.

The database also provides a range of demographic information for each investor including the investor's age, gender, language and postcode of residence. We use these variables as controls and to help confirm the validity of our social interaction proxy. In order to calculate distances between the residences of investors, we obtain values for the latitude and longitude of postcode areas from ZIPCodeSoft. In addition, every account is associated with an ID number representing the individual investor's family name. Investors with the same family name receive the same ID number. We use this variable as a proxy for familial relationships. Finally, we obtain stock price and firm fundamentals data from COMPUSTAT Global.

2.2 Identification of social interaction

In order to measure social interaction between investors, we apply the techniques proposed by Ozsoylev et al. (2014) to identify "Empirical Investor Networks". As discussed below, we introduce some modifications to allow a more accurate identification of interaction. The major advantage of this approach is that connections can be identified between the entire population of investors, as long as account-level trading data is available. In section 3 of our paper, we provide strong evidence that our investor network reflects true social interaction between investors.

According to Ozsoylev et al.'s (2014) methodology, two investors are defined as being connected if they trade the same stock, s, in the same direction within a defined time period, Δt, a given number of times, m, during a sample window of length l. The motivation behind this approach is that investors who communicate with each other in relation to trading decisions are likely to trade the same stocks within a short space of time. Given the context of our analysis, we do not include sell trades in identifying connections for two reasons. First, buy transactions reflect a much more active choice by investors than sell transactions and better represent an investor's current attitudes in relation to the riskiness of the stock they wish to acquire. Investors are much more limited in their sell decisions (they must first own the stock) and the stocks that do own are likely to reflect prior rather than current attitudes towards risk. Second, certain sell transactions are likely to be primarily determined by liquidity needs.[2] In any given year, we define as active and include investors who make buy trades on at least ten days during the year.

Estimation of the investor network requires choices for each of the variables described above. Like Ozsoylev et al. (2014), we work with a sample window of one year. That is, when conducting time-series tests across a number of years, network connections are recalculated each year. We experiment with different values for Δt and m but the majority of the results we present use Δt equal to one day and m equal to 5. There is a trade-off between specifying lower and higher values for Δt and m respectively to increase our level of confidence that two individuals truly are connected and designating two investors as unconnected when they are actually connected. We believe that a time period equal to at least one day is appropriate for the type of word-of-mouth interaction between retail investors that we are attempting to capture. As described below, we do not include high-frequency traders in our sample and a time period of less than one day would be too restrictive. Furthermore, we also wish to capture instances where two individuals interact about a particular stock on a given day and then both trade the stock the following day. In summary, therefore, two investors are defined as being connected during a calendar year if, on at least 5 days during the year, they purchased the same stock.

2.3 Localised social interaction versus reactions to centralised information diffusion

A potential issue related to the identification of Empirical Investor Networks purely from common trades is that the trades may reflect reactions to public diffusion of important

[2] However, our results are robust to an alternative specification which includes both buy and sell trades in identifying connections.

news events or common trading strategies rather than true interaction between investors. We provide an important contribution to this area of the literature by evaluating the extent to which these factors may bias the identification of an Empirical Investor Network. We first identify all the connected buy trades for each stock: instances where two investors purchase the same stock on the same day. For each stock which has at least 100 connected buy trades during the year, we then compute the percentage of the total yearly connected buy trades that occur on each trading day. We then rank the trading days within each stock-year and compute the average for each ranked trading day. As shown in Figure 1, the distribution shows extreme right skewness. The highest ranked trading day accounts for more than 20% of connected trades on average. Together, the top five ranked trading days contain an average of about 45% of connected trades. This means that for each stock, a small number of trading days are responsible for a disproportionately large share of connected trades.

In order to determine whether these trading days correspond to the release of important stock-specific news which would be expected to induce high trading volume, we identify a total of 4,781 earnings announcement days between 2000 and 2010. We find that the average percentage of yearly connected buy trades which occur on earnings announcement days is 2.34%, higher by a factor of 12 relative to the 0.19% which occur on non-earnings announcement days. These results heighten the concern that a significant portion of the connected trades we identify might be reactions to centralised diffusion of news events rather than the localised social interaction between specific investors which we intend to capture. We impose four additional controls in order to address this issue.

First, when specifying the network, we ignore days on which a high amount of public information-induced trading is likely. Because we cannot directly identify all the potential sources of public information diffusion, we proxy for such days based on our analysis above. Specifically, we ignore connections occurring on days which account for more than 1% of total annual connected trades for a given stock. In other words, to be identified as socially connected, two investors must buy together on days *other* than high information days. Second, we eliminate investors who are identified as being connected to more than five other individuals. Therefore, an investor’s trades must be shared with only a small number of network peers, something which is unlikely to occur if a large number of investors are responding to public information diffusion.

Third, we eliminate high-frequency investors whom we define as trading on more than 100 days during a given year. High-frequency traders who trade several stocks daily are quite likely to be identified as being connected to other high-frequency traders simply due to the trading strategies they follow rather than any actual communication occurring between them. Moreover, our aim in this study is to examine the effects of social interaction between ordinary retail investors rather than high-frequency or day traders. We find that very few investors trade more than 100 times during the year (about 2% per year). Finally, in order to help avoid identifying investor pairs where one member of the pair trades on behalf of the other member, we limit our sample to investors who are aged at least 25. For example, Berkman, Koch and Westerholm (2013) document evidence consistent with Finnish parents and guardians trading on behalf of their children’s accounts.

2.4 Advantages of the identification methodology

Our method of identifying social interaction brings a number of advantages. Firstly, it enables us to classify any pair of investors as being either socially connected or unconnected based solely on trade-level data. In addition, coupled with the validation tests we perform below, it gives us confidence that we are capturing interaction between investors which is directly related to the investment decision-making process. Most prior empirical papers have relied on more general and aggregate measures to proxy for social interaction including geographical proximity (Ivković & Weisbenner, 2007), church attendance (Hong, Kubik, & Stein, 2004) or participation in sports events (Heimer, 2014). Such measures are better suited to an aggregate setting as they do not allow social connections between specific individuals to be identified. Moreover, the type of social interaction they identify may not necessarily translate to interaction about investments.

2.5 Event study design

We conduct our analysis in the form of an event study. During each year of our sample (1997 to 2011) we construct the social network of investors. For a given investor, we define year zero as the first year during which we identify the investor as being socially connected to at least one other investor. As describe above, two investors are connected if, on five or more occasions during the year, they purchase the same stock. Given that our minimum five shared buy trades cut-off is a relatively stringent restriction, it is possible that some of our connected pairs were already interacting prior to year zero, which would add noise to the sample. To help ensure that this is not the case, we retain only those pairs of individuals who made no common buy trades at all prior to year zero. In order to be able to track an investor's behaviour over time, we require that the investor be active during at least three years prior to year zero. Table 1 provides summary statistics for the investors defined as being socially interactive during each event year from year -4 through year +4. During year zero, 1,783 investors are contained in our sample. A relatively constant proportion of between 10% and 17% of the sample are female and the mean age increases from 51 in year -4 to 57 in year +4. The median trading frequency is highest during the year of first connection, suggesting that investors are particularly active during this year.

The summary statistics suggest that some variation exists in the characteristics of investors making up the sample each event year. An important advantage of our event study setting is that it allows us to track variation in the behaviour of the same investor over time as he or she becomes socially interactive, thus allowing us to implicitly control for innate and time-invariant individual characteristics. We explicitly control for time-variant characteristics such as trading frequency and trade size in multivariate regressions below. Moreover, the fact that year zero can correspond to a different calendar years for different investors helps control for time period specific effects which might influence the trading behaviour of all investors in the sample.

2.6 Trading risk measure

In order to test our two competing hypotheses, we require a measure for an investor's tendency to follow high-risk or extreme trading strategies. A very direct measure of this type of trading behaviour is the proportion of an investor's trades in high-volatility and high-skewness stocks. Both of these stock characteristics are specifically mentioned by Han and Hirshleifer (2012) as reflecting the type of active trading promoted through social interaction. For each calendar year in our sample, we sort stocks into quintiles based on volatility, idiosyncratic volatility and skewness. We calculate a stock's annual volatility and skewness using the stock's daily returns during the year. Idiosyncratic volatility is calculated as the standard deviation of residuals from a regression of daily stock returns on the market index during the year. We only include stocks with at least 100 daily return observations during the year.

For each investor in our sample, we then compute the proportion of total trading value during the year that is accounted for by stocks falling into the highest quintile for a given stock characteristic. For example, we denote $PROP\_HVOL_{i,t}$ as the proportion of investor i's total trading value in year t represented by stocks in the highest volatility quintile.

In order to provide a clearer measure of the proportion of trades by an individual in high-risk stocks in excess of what we would expect, we construct an adjusted, or "abnormal", proportion, which we term APROP. To do this, we subtract the average proportion that other investors in the market with a similar trading frequency invest in high-risk stocks. Specifically, we divide all investors in the market into groups based on their trading frequency and calculate market-wide proportions as the average within each frequency group.[3] For example, we calculate $APROP\_HVOL_{i,t}$, the abnormal proportion of investor i's total trading value in year t represented by stocks in the highest volatility quintile, as follows:

[3] Our groups are formed based on the frequency of trades in lots of 10. For example, the first group consists of investors with a trading frequency between 10 and 20, the second of investors with a trading frequency between 21 and 30 and so on.

$$APROP\_HVOL_{i,t} = PROP\_HVOL_{i,t} - \overline{PROP\_HVOL}_t. \quad (1)$$

Here, $\overline{PROP\_HVOL}_t$ is the average proportion of trading value in year t represented by stocks in the highest volatility quintile across all investors (excluding investor i) in the same frequency group as investor i.

To this end, we divide investors into two groups based on their observed risk-taking during the period preceding the year in which they are first identified as socially connected. Specifically, we designate the investors who display an average value of APROP less than zero in the pre-period as "conservative" investors and those that display an average value above zero as "speculative" investors. We find that investors in our sample are approximately evenly split between the two groups, with a slight tilt towards the conservative style. Between 55% and 57% of investors are classed as conservative using the three different measures of risk.

Moreover, we also track the trading behaviour of an individual's social network during each event year. Individual i's social network consists of the investor (or investors) with which he becomes socially connected in year zero. In most cases, individuals are identified as interacting with only one other investor in year zero. In cases where an individual's social group consists of more than one other investor (the maximum is five), we take the average trading behavior across the investors in the social group.

3 Validity of social interaction measure

Our ability to infer a relationship between social interaction and subsequent trading behaviour depends on the validity of our measure of social interaction. We verify the validity of our approach by testing whether the connections we identify tend to be clustered around individuals whom we would expect to be connected due to observable exogenous traits. Such traits include close geographical proximity, familial relationships and common demographic characteristics.

Our analysis in this section is also interesting in its own right. In constructing their empirical investor network, Ozsoylev et al. (2014) have no information about the characteristics of the investors in their sample. This leaves a number of open questions pertaining to the nature of the connections they identify. By observing such characteristics as geographical location, gender, age and familial relationships, we contribute to our understanding of the channels by which information diffuses within an investor network.

3.1 Relation between social interaction and investor characteristics

We test whether the propensity of two investors to be connected in the network is related to geographical proximity between the two investors and shared personal characteristics. We base our measure of geographical proximity on the distance (in km) between the coordinates of the postcode areas in which the investors reside calculated using the Vincenty formula. The small size of Finland's postcode areas means that we obtain a large amount of variation in distances between investors. There are over 3,000 distinct postcodes in our sample. We follow Pool, Stoffman and Yonker (2013) by normalising the distance to account for variations in population density in different areas. We refer to our adjusted distance measures as "normalised-equivalent distances". The same number of people live within a given normalised-equivalent distance of each other as live within the corresponding actual distance of each other.[4] A full description of the adjustment is given in Appendix A. The additional investor characteristics we consider are familial relationships, gender, age and language. We present results for 2011, the most recent year in our sample period.

The first two rows in Table 2 show the number of potential and actual connections that we observe during 2011. For example, during 2011, there were a total of 25,756 investors meeting our restrictions in the sample. The number of potential connections can be calculated as $\frac{25{,}756^2 - 25{,}756}{2}$ = 331,672,890. The number of actual connections that we observe during the year is 681. In order to evaluate how the distribution of these connections is related to investor characteristics, we next compare the expected and actual percentage of connections falling within different subgroups of investors. Panel A groups investors according to geographical proximity while Panel B uses other demographic traits.

The expected percentage of connected pairs is calculated according to the null hypothesis that the propensity that two investors are connected is independent of their geographical proximity and respective demographic characteristics. Specifically, the expected proportion of investor pairs that are connected for a particular subgroup of investors is calculated as follows:

$$Expected\ proportion = \frac{Potential\ connections\ within\ subgroup}{Total\ potential\ connections}. \qquad (2)$$

[4] For example, the same number of people lives within 5 normalised equivalent km of each other as lives within 5 actual kilometres of each other. However, the specific individuals will not necessarily be the same in each group.

For example, the potential number of connections between two investors whose postcodes of residence are no more than 1 normalised-equivalent km apart is 0.763 million. This represents 0.763M/331.672M = 0.23% of total potential connections in the network. However, as shown in the following row of Table 2, the observed proportion of connections that are represented by two investors whose postcodes of residence are no more than 1 normalised-equivalent km apart is 138/681 = 20.26%. This percentage is higher than the expected value by a factor of 88, suggesting that two investors are much more likely to be connected if their postcodes of residence are no more than 1 normalised-equivalent km apart relative to the population of investors as a whole. Z-statistics and p-values are provided for the null hypothesis that the actual proportion is equal to the expected proportion. As expected, the ratio of the actual percentage to the expected percentage falls the higher the normalised-equivalent distance between two individuals.

Figure 2 provides a visual representation of the relationship between geographical proximity and social interaction. We plot the expected and actual cumulative distribution functions (CDFs) of distances between residences of socially connected investors. To aid interpretation, we use true rather than normalised-equivalent distances for the figure. Relative to the expected CDF, the actual CDF initially increases as a much steeper gradient. 50% of the social connections we observe occur at a distance of less than 80km; according to the null hypothesis that the probability of being connected is independent of proximity, we would expect only 20%. These results provide strong evidence in favour of our measure of social interaction. Individuals who live close to each other are much more likely to be connected than those who live far apart.

We assess the relation between the propensity of being identified as connected and other demographic traits in Panel B. The results reveal that familial relationships (proxied by common surnames) are a particularly strong predictor of connections. Just over 20% of the connections we identify occur between individuals who have the same family name. This is higher than the proportion predicted by the null hypothesis by a factor of 364. This result is also interesting because it suggests that family relationships may act as a significant influence on investment decisions, a finding which has received little attention in the literature. We find that two individuals are slightly less likely to be connected than expected if they are of the same gender. This likely reflects the impact of partner relationships. Finally, investors are more likely to be identified as connected than expected under the null if they speak the same language and if they are of similar ages.

Taken together, the results from Table 2 and Figure 2 present a strong case that our investor network is proxying for the types of word-of-mouth interaction effects between investors that we intend to pick up. Channels of information diffusion are particularly strong among investors who reside in close geographical proximity, have familial relationships and are otherwise similar in terms of age and language. A priori, we would expect such investors to have a much greater chance of knowing and interacting with each other.

4 Results

Having established the validity of our proxy for social interaction, we now move on to examining the relationship, if any, between social interaction and the risky trading behaviour displayed by individuals. To recap, we are mainly interested in differentiating between two competing hypotheses about the impact of social interaction on behaviour. On the one hand, social interaction has been suggested as a catalyst for intensifying the prevalence of high-risk trading strategies and behavioural biases. On the other hand, social interaction could act as a moderating factor on extreme strategies followed by individuals and encourage more rational decision making.

4.1 Social interaction as a moderator of behaviour

4.1.1 Univariate tests

We begin by tracking how the risk-related trading behaviour of conservative and speculative investors changes over time as they are first identified as socially connected. A visual representation of the results is presented in the charts in Figure 3. Conservative and speculative investors are depicted in Panels A and B respectively. The black line tracks APROP for investor i in event time while the red circles track APROP for investor i's social network. Charts are plotted separately for the three risk measures. The results are strongly supportive of social networks acting as a moderating influence on individual investor behaviour. The charts show a strong movement of an individual's value of APROP away from the relatively extreme values and towards the more average trading behaviour of his network peers. Conservative investors tend to increase their level of risk exposure while speculative investors decrease their exposure. In addition, the steepest change in risk exposure occurs during the year directly prior to the year in which investors are first identified as connected. However, the charts also indicate an asymmetry between conservative and speculative individuals. Speculators appear to be much more subject to the moderating effect than conservatives. It is interesting to note that this pattern is consistent with the finding by Masclet, Colombier, Denant-Boemont and Loheac (2009) that less risk averse individuals show a greater tendency to shift their decision to conform to the group average in an experimental setting.

In order to examine the statistical significance of these effects, we calculate the mean value of APROP in the pre-connection and post-connection (including year zero) periods for each individual. We also take the difference between the post- and pre-connection means. We then calculate averages and associated t-statistics separately across conservatives and speculators. The results, presented in Table 3, show that the patterns of risk-taking behaviour shown in the charts are statistically significant. For example, using the volatility measure of risk, conservative investors display a statistically significant increase in APROP from -0.08 in the pre-connection period to -0.03 in the post period. For higher-risk individuals, the change is even more dramatic, falling from 0.12 in the pre-connection period to 0.04 in the post-connection period. Similar results are obtained when we use idiosyncratic volatility or skewness as our measure of risk.

4.1.2 Regressions: Levels

In order to provide a more robust estimate of the relation between the risky trading behaviour of an individual and that of his social network, we use a regression approach. Our regression equation is similar to the one applied by Ahern, Duchin, & Shumway (2014) to identify the impact of social interaction on risk aversion and trust in a longitudinal sample of MBA students. Specifically, we estimate the following model:

$$APROP_{i,0} = \alpha + \beta_1 \overline{APROP\_PRE\_NW}_i + \gamma X_{i,0} + \varepsilon_{i,0}. \tag{3}$$

Here, the dependent variable is the value of APROP for individual i during the year of first connection.[5] The main explanatory variable of interest is $\overline{APROP\_PRE\_NW_i}$, which refers to the average value of APROP across the individuals comprising investor i's social network during the pre-connection period. A convergence of individual i's trading behaviour towards that of his or her social peers implies a positive coefficient β_1. Additional control variables include age, gender, trading frequency and the log of an investor's average trade value. It is important to note that the trading behaviour of an individual's network is measured in the pre-connection period only. This avoids the mechanical positive relationship which would obviously result from the fact that our method of identifying social connections is based on observing common trading behaviour.

Following the approach in the univariate analysis above, we also introduce a dummy variable equal to one for investors designated as speculators and zero otherwise. We interact this dummy with the pre-period average value of APROP for an investor's network as follows:

$$\begin{aligned} APROP_{i,0} = \alpha + \beta_1 \overline{APROP\_PRE\_NW_i} + \beta_2 SPECULATOR_i + \beta_3 SPECULATOR_i \\ \times \overline{APROP\_PRE\_NW_i} + \gamma X_{i,0} + \varepsilon_{i,0}. \end{aligned} \tag{4}$$

The coefficient on the interaction term, β_3, reflects whether or not the convergence of trading behaviour between an individual and his network is stronger for those individuals who tend to display above-average risk in the pre-connection period. A positive coefficient indicates stronger convergence for speculative investors while a negative coefficient indicates stronger convergence for conservative individuals.

---

[5] We obtain similar results if we use the average APROP over the post-connection years as the dependent variable instead.

Table 4 displays the coefficient estimates for the two regression specifications for each of the three measures of risk. Looking first at the results for equation 3 (model numbers 1, 3 and 5 in Table 4), the coefficient estimate β_1 is positive and statistically significant at higher than the 1% level in all cases. The magnitudes of β_1 also appear to be significant in economic terms. For example, an increase by one percentage point in the abnormal proportion that an individual's social network trades in high-volatility stocks during the pre-connection period is associated with an increase of 0.29 percentage points in the individual's own abnormal portion traded in high-volatility stocks. The coefficient estimate is similar when APROP is measured using idiosyncratic volatility or skewness.

When we estimate regression equation 4 and differentiate between lower and higher risk investors (model specifications 2, 4 and 6 in Table 4), the asymmetry of the moderating effect is again apparent. Significantly positive estimates of β_1 when APROP is defined using volatility or idiosyncratic volatility shows that a moderating effect of social interaction operates for conservative investors. An increase by one percentage point in the abnormal proportion that a conservative individual's social network trades in high-volatility stocks during the pre-connection period is associated with an increase of 0.09 percentage points in the individual's own abnormal portion traded in high-volatility stocks. However, the coefficient on the interaction term, β_3, shows that the moderating effect is considerably stronger for speculative individuals. For example, an increase by one percentage point in the abnormal proportion that a speculator's social network trades in high-volatility stocks during the pre-connection period is associated with an increase of 0.43 percentage points (0.09 + 0.34) in the individual's abnormal portion traded in high-volatility stocks. This asymmetry is equally apparent when idiosyncratic volatility or skewness are used to define APROP. In the case of skewness, the convergence effect appears to be limited to speculative individuals, as shown by an insignificant coefficient β_1.

Significantly positive estimates of β_2, the coefficient on the speculative investor dummy variable indicate that the risk-related trading behaviour that investors exhibited during the pre-connection period does persist to some extent during the year of first connection. Also, the abnormal portion of trades invested in high risk stocks tends to be negatively related to average trade value and positively related to trading frequency. There is little suggestion of any relationship with either gender or age.

4.1.3 Regressions: Changes

An alternative way of measuring the relation between investors' risky trading behaviour and that of their network peers is by focusing on changes in behaviour. We regress the change in APROP for a given investor from the pre-connection period average to the year of first connection on the difference between the average levels of APROP for the investor and that of his network during the pre-connection period. Specifically, we estimate:

$$
\begin{aligned}
APROP_{i,0} &- \overline{APROP\_PRE_i} \\
&= \alpha + \beta_1(\overline{APROP\_PRE_i} - \overline{APROP\_PRE\_NW_i}) + \gamma X_{i,0} + \varepsilon_i.
\end{aligned} \tag{5}
$$

In this case, a convergence of investors' risk-related trading behaviour towards that of their network peers would be indicated by a significantly negative coefficient β_1. This would indicate that a larger deviation between the abnormal risk exposures between investors and their peers during the pre-interaction period is associated with a larger shift in risk exposure (in the opposite direction) as investors become connected. Again, we also include a specification with a dummy variable equal to one for speculative investors and zero otherwise. We interact this dummy with the pre-period average difference in APROP between the investors and their network peers to detect any asymmetry in the convergence between the conservative and speculative investor groups.

The results, presented in Table 5, show additional evidence of a convergence of the risky trading behaviour of investors to that of their network peers. The coefficient estimate, β_1, for regression equation 5 (shown in columns 1, 3 and 5 of Table 5) is negative and statistically significant for each of the three risk measures. This indicates that the larger the difference between the portion that individuals and their network peers invest in high-risk stocks during the pre-connection period, the greater the decrease in the proportion that an individual invests in such stocks once he becomes connected. For the case where APROP is measured using volatility, an increase of 100 basis points in the pre-period difference is associated with a 28 basis point decrease in an investor's level of APROP between the pre-period average and year zero. The asymmetry of this effect between conservative and speculative investors is again demonstrated by the significantly negative coefficients on the interaction term (columns 2, 4 and 6 in Table 5). The results hold for all three measures of APROP.

4.2 Aggregate effect on risk-taking

The tests described above document a convergence in risk-related trading behaviour by both conservative and speculative investors towards the market norm. In this section, we briefly consider the aggregate relation between social interaction and risk taking for the market as a whole. The notion that social interaction induces purchasing of high-risk stocks would be reflected by an overall rise in levels of APROP following the year of first interaction. In contrast, given the observation that the moderating effect of social interaction appears stronger for speculative than conservative investors, we expect the proportion of funds invested in high-risk stocks for investors overall to fall when investors first begin to interact. In order to test this, we run investor-level regressions in event time. For each of our three risk measures, we estimate the following regression equation:

$$\begin{aligned} APROP_{i,t} = \propto + \beta_1 POST_t + \beta_3 TRADING\_FREQ_{i,t} \\ + \text{INVESTOR FIXED EFFECTS } + \varepsilon_{i,t} \end{aligned} \tag{6}$$

The model includes a dummy variable, $POST_t$, which is equal to 1 from the year of first connection onwards. We control for an investor's trading frequency with a separate control variable and any time-invariant investor characteristics such as age, gender or personal traits are controlled for via investor fixed effects. Table 6 displays the results. The coefficient estimate on our main variable of interest, the dummy for the year of first connection, is negative and statistically significant at the 5% level or higher when APROP is defined using idiosyncratic volatility or skewness. When APROP is defined using volatility, the coefficient is also negative but insignificant. Overall, the results are consistent with our previous tests and suggest that investors tend to reduce the proportion of trades allocated to high-risk stocks when they become socially connected. These results do not appear to support the notion that social interaction enhances high-risk trading strategies as predicted by Han and Hirshleifer (2012).

4.3 Performance

A commonly cited notion in relation to retail investor behaviour is that individuals tend to trade too actively and over-invest in high-risk and high-skewness stocks. To the extent that this leads to underdiversification and exposure to overvalued securities, portfolio performance can suffer. Thus, the idea that social interaction is associated with risky trading behaviour suggests that it might also impact portfolio performance. In this section, we briefly examine this issue by following the alphas earned on portfolios formed based on the buy decisions of investors displaying different levels of risk taking. This time, we divide investors into three risk groups based on their quintile ranking of APROP during the pre-connection period. Risk group 1 contains investors falling into quintile 1, risk group 2 contains those falling into quintiles 2-4 and risk group 3 contains those falling into quintile 4.

We estimate alphas using a calendar time portfolio approach. First, on each trading day, we further divide the investors contained in each risk group into a connected and an unconnected subgroup. Investors move into the connected subgroup on the first day of the year in which they are first identified as connected. For each group and subgroup of investors, we form a portfolio based on the investors' prior stock purchase decisions. On each trading day, the portfolio takes a long position in the stocks for which the investors in the group who were net buyers of the stock during the last 90 calendar days outnumber those who were net sellers. We then regress the daily return of this equally-weighted portfolio against the three Fama French factors to estimate the risk-adjusted daily alpha.

Table 7 displays the results for the different risk groups and the unconnected and connected subgroups. Panels A, B and C use APROP calculated using volatility, idiosyncratic volatility and skewness respectively. In the majority of specifications, estimated alphas range between 0.02% and 0.03% with varying degrees of statistical significance. There is no evidence that trading performance, as measured by the buy portfolio alpha, suffers following the initiation of social interaction. Out of a total of nine specifications (three risk groups × three APROP measures), the alpha is lower for the connected subgroup relative to the unconnected subgroup in only two cases. For both of these cases, the differences are not statistically significant and relate to the lowest risk group. In the remaining seven specifications, the value is higher for the connected relative to the unconnected subgroup. The differences appear to be largest for individuals who display the highest levels of APROP during the pre-connection period (risk-group 3). Again, the differences are not statistically significant however. Overall, therefore, there is little evidence that portfolio performance differs significantly before and after an individual becomes socially connected. Importantly, the notion that performance suffers following social interaction is unsupported.

5 Conclusion

A recent theory by Han and Hirshleifer (2012) proposes that social interaction plays a role as a catalyst for active and high-risk investment strategies. This contrasts with implications from the experimental economics literature, in which group interaction has been found to act as a moderating influence on behaviour. This paper is the first to empirically test these conflicting predictions in a retail stock market investor context. We find evidence consistent with the idea that interaction has a moderating effect on risk taking. Moreover, the effect is stronger for investors who display above-median levels of risk taking prior to interacting for the first time.

A recent paper by Ahern et al., (2014) finds that the level of risk aversion displayed by MBA students is strongly influenced by peer effects, implying that risk aversion is transitory and easily influenced by environmental settings. In our study, we observe only risk outcomes and therefore cannot directly relate our findings to underlying risk attitudes. Nevertheless, our results suggest that risk-related investment behaviour is easily subject to external influence, which seems in line with the findings by Ahern et al., (2014).

The notion that social interaction may act as a moderating influence on more extreme investor behaviour is new to the investment literature and could have some important implications. Social interaction may actually be of benefit to investors with a tendency to overinvest in high-risk and high-skewness stocks if this tendency is associated with underperformance.

Appendix A

In testing the validity of our proxy for social interaction, we require a measure of geographical proximity between two investors i and j. We base this measure on the distance (in km) between the coordinates of the postcode areas in which the investors reside calculated using the Vincenty formula. We follow Pool, Stoffman and Yonker (2013) by normalising the distance to account for variations in population density in different areas. The normalised distance between investors i and j is calculated as follows:

$$NORM\_DIST_{i,j} = DIST_{i,j} \times \frac{\mathrm{MAX}(POPULATION\_DENSITY_i, POPULATION\_DENSITY_j)}{MEDIAN(POPULATION\_DENSITY)}. \qquad (1)$$

The population density is taken from the Statistics Finland website and corresponds to the area in which an investor resides. The median is calculated across all the investors in the sample. What we refer to in the paper as the "normalised-equivalent distance" corresponds to the normalised distance that contains the same number of investors as the stated standard distance measure. For example, the same number of investors live "within 50 normalised-equivalent km" of each other as live within 50 km (in standard terms) of each other. The actual individuals in the two groups will not be the same however since the normalised version takes population density into account. The values for the normalised km corresponding to normalised-equivalent distances of 1, 5, 10, 20 and 50km respectively are 0.32, 5.09, 12.03, 26.99 and 65.65.

References

Ahern, K. R. (2014). Information networks: Evidence from illegal insider trading tips. *Available at SSRN,*

Ahern, K. R., Duchin, R., & Shumway, T. (2014). Peer effects in risk aversion and trust. *Review of Financial Studies,* , hhu042.

Baghestanian, S., Gortner, P. J., & Van der Weele, Joel J. (2014). Peer effects and risk sharing in experimental asset markets.

Banerjee, A. V. (1992). A simple model of herd behavior. *The Quarterly Journal of Economics, 107*(3), 797-817.

Banerjee, A., & Fudenberg, D. (2004). Word-of-mouth learning. *Games and Economic Behavior, 46*(1), 1-22.

Berkman, H., Koch, P. D., & Westerholm, P. J. (2013). Informed trading through the accounts of children. *The Journal of Finance,*

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy,* , 992-1026.

Brown, J. R., Ivković, Z., Smith, P. A., & Weisbenner, S. (2008). Neighbors matter: Causal community effects and stock market participation. *The Journal of Finance, 63*(3), 1509-1531.

Cao, H. H., Han, B., & Hirshleifer, D. (2011). Taking the road less traveled by: Does conversation eradicate pernicious cascades? *Journal of Economic Theory, 146*(4), 1418-1436.

Cheung, S. L., & Palan, S. (2012). Two heads are less bubbly than one: Team decision-making in an experimental asset market. *Experimental Economics, 15*(3), 373-397.

Duflo, E., & Saez, E. (2002). Participation and investment decisions in a retirement plan: The influence of colleagues' choices. *Journal of Public Economics, 85*(1), 121-148.

Ellison, G., & Fudenberg, D. (1993). Rules of thumb for social learning. *Journal of Political Economy,* , 612-643.

Ellison, G., & Fudenberg, D. (1995). Word-of-mouth communication and social learning. *The Quarterly Journal of Economics, 110*(1), 93-125.

Guiso, L., & Jappelli, T. (2005). Awareness and stock market participation. *Review of Finance, 9*(4), 537-567.

Han, B., & Hirshleifer, D. A. (2012). Self-enhancing transmission bias and active investing. *Available at SSRN 2032697,*

Heimer, R. Z. (2014). Friends do let friends buy stocks actively. *Journal of Economic Behavior & Organization,*

Hellström, J., Zetterdahl, E., & Hanes, N. (2013). Loved ones matter: Family effects and stock market participation.

Hong, H., Kubik, J. D., & Stein, J. C. (2004). Social interaction and stock-market participation. *The Journal of Finance, 59*(1), 137-163.

Hong, H., Kubik, J. D., & Stein, J. C. (2005). Thy neighbor's portfolio: Word-of-Mouth effects in the holdings and trades of money managers. *The Journal of Finance, 60*(6), 2801-2824.

Horton, J., & Serafeim, G. (2009). Security analyst networks, performance and career outcomes.

Hvide, H. K., & Östberg, P. (2013). Stock investments at work.

Ivković, Z., & Weisbenner, S. (2007). Information diffusion effects in individual investors' common stock purchases: Covet thy neighbors' investment choices. *Review of Financial Studies, 20*(4), 1327-1357.

Kaustia, M., & Knüpfer, S. (2012). Peer performance and stock market entry. *Journal of Financial Economics, 104*(2), 321-338.

Kugler, T., Kausel, E. E., & Kocher, M. G. (2012). Are groups more rational than individuals? A review of interactive decision making in groups. *Wiley Interdisciplinary Reviews: Cognitive Science, 3*(4), 471-482.

Liu, Y., Meng, J., You, W., & Zhao, L. (2013). Word-of-mouth communication, observational learning, and stock market participation. *Observational Learning, and Stock Market Participation (April 15, 2013),*

Lu, T. J. (2012). Social interaction effects and individual portfolio choice: Evidence from 401 (k) pension plan investors. *Available at SSRN 2024095,*

Masclet, D., Colombier, N., Denant-Boemont, L., & Loheac, Y. (2009). Group and individual risk preferences: A lottery-choice experiment with self-employed and salaried workers. *Journal of Economic Behavior & Organization, 70*(3), 470-484.

Ozsoylev, H. N., Walden, J., Yavuz, M. D., & Bildik, R. (2014). Investor networks in the stock market. *Review of Financial Studies, 27*(5), 1323-1366.

Pool, V. K., Stoffman, N., & Yonker, S. E. (2013). The people in your neighborhood: Social interactions and mutual fund portfolios. *Journal of Finance, Forthcoming,*

Rockenbach, B., Sadrieh, A., & Mathauschek, B. (2007). Teams take the better risks. *Journal of Economic Behavior & Organization*, *63*(3), 412-422.

Shive, S. (2010). An epidemic model of investor behavior. *Journal of Financial and Quantitative Analysis, 45*(1), 169.

Simon, D., & Heimer, R. (2012). Facebook finance: How social interaction propagates active investing. *AFA 2013 San Diego Meetings Paper,*

Welch, I. (1992). Sequential sales, learning, and cascades. *The Journal of Finance, 47*(2), 695-732.

Table 1: Summary statistics by event year

This table displays summary statistics for the investors active during each event year between year -4 and year +4. Event year zero refers to the year during which we first classify an investor as being socially connected. Two investors are defined as being connected during a calendar year if, on at least 5 days during the year, they purchased the same stock. In order to be included in the sample, investors must be active during at least three years prior to year zero. Investors who make fewer than 10 trades during a calendar year are not considered to be active during that year and investors who trade more than 100 times during any calendar year are excluded from the sample. The sample period is 1997 to 2011.

	Event year								
	-4	-3	-2	-1	0	1	2	3	4
N	977	1,251	1,286	1,394	1,783	1,109	674	403	242
Percent female	11	11	10	11	11	11	12	12	17
Mean age	51	51	51	53	54	55	55	57	57
Median trading frequency	18	18	21	22	35	29	26	28	25

Table 2: Distribution of connections by geographic proximity and demographic characteristics

This table shows how connections between pairs of investors during 2011 are distributed according to geographical proximity between the two investors and shared demographic characteristics. Two investors are defined as being connected during a calendar year if, on at least 5 days during the year, they purchased the same stock. The table compares the expected and actual proportions of the total connections observed in the network that are represented by investors falling into different subgroups. The expected proportion is calculated according to the null hypothesis that the propensity that two investors are connected is independent of their geographical proximity and respective demographic characteristics: $Expected\ proportion = \frac{Potential\ connections\ within\ subgroup}{Total\ potential\ connections}$. Panel A forms subgroups based on geographic proximity while Panel B uses demographic characteristics. The Normalised-equivalent distance adjusts the actual distance for population density as described in the text. Z-statistics and p-values are provided for the null hypothesis that the actual proportion is equal to the expected proportion.

Panel A: By geographical proximity	Entire network	Normalised-equivalent distance(1,2) <= 1km	Normalised-equivalent distance(1,2) <= 5km	Normalised-equivalent distance(1,2) <= 10km	Normalised-equivalent distance(1,2) <= 20km	Normalised-equivalent distance(1,2) <= 50km
Potential connections	331,672,890	763,205	6,202,155	16,915,259	33,731,132	52,599,918
Actual connections	681	138	174	208	247	279
Expected %		0.23	1.87	5.10	10.17	15.86
Actual %		20.26	25.55	30.54	36.27	40.97
Actual %/Expected %		88.06	13.66	5.99	3.57	2.58
z		109.140	45.620	30.180	22.530	17.940
p		<0.001	<0.001	<0.001	<0.001	<0.001

Panel B: By demographic factors	Entire network	Same family	Same gender	Same language	Age diff <= 5y	Age diff <= 10y
Potential connections	331,672,890	198,186	247,136,295	295,255,921	72,854,848	133,487,174
Actual connections	681	148	423	624	253	370
Expected %		0.06	74.51	89.02	21.97	40.25
Actual %		21.73	62.11	91.63	37.15	54.33
Actual %/Expected %		363.71	0.83	1.03	1.69	1.35
z		230.960	-7.420	2.180	9.572	7.495
p		<0.001	<0.001	0.029	<0.001	<0.001

Table 3: Moderating effects of social interaction for conservative and speculative investors

This table displays descriptive statistics for the average value of APROP in the pre-connection and post-connection (including year zero) periods across investors as well as the post- and pre-connection means. APROP refers to the abnormal proportion of an investor's total trade value accounted for by stocks in the highest risk quintile during a given year where risk is measured by volatility, idiosyncratic volatility and skewness (shown in Panels A through C respectively). The connection year (year zero) refers to the year during which we first classify an investor as being socially connected. Two investors are defined as being connected during a calendar year if, on at least 5 days during the year, they purchased the same stock. Results are shown separately for investors who display a below- (conservative) and above-median (speculative) average value of APROP during the pre-connection period. The sample period is 1997 to 2011.

Panel A: APROP measured by volatility

	Conservative investors			Speculative investors		
	N	Mean	t	N	Mean	t
Pre-connection	1,017	-0.08	-45.55	766	0.12	28.72
Post-connection	1,017	-0.03	-8.17	766	0.04	6.15
Post-connection - Pre-connection	1,017	0.05	13.73	766	-0.08	-13.32

Panel B: APROP measured by idiosyncratic volatility

	Conservative investors			Speculative investors		
	N	Mean	t	N	Mean	t
Pre-connection	1,033	-0.07	-45.63	750	0.11	27.73
Post-connection	1,033	-0.02	-10.97	750	0.03	5.92
Post-connection - Pre-connection	1,033	0.04	19.37	750	-0.08	-15.16

Panel C: APROP measured by skewness

	Conservative investors			Speculative investors		
	N	Mean	t	N	Mean	t
Pre-connection	995	-0.05	-49.97	788	0.08	27.58
Post-connection	995	-0.02	-6.61	788	0.02	4.55
Post-connection - Pre-connection	995	0.03	13.31	788	-0.06	-13.54

Table 4: Regression – Moderating effects of social interaction for conservative and speculative investors (Levels)

This table displays coefficient estimates for OLS regressions where the dependent variable is an investor's APROP during event year zero. APROP refers to the abnormal proportion of an investor's total trade value accounted for by stocks in the highest risk quintile where risk is measured by volatility, idiosyncratic volatility and skewness. Pre-connection average network APROP refers to the average value of APROP across the individuals comprising the investor's social network during the pre-connection period. The speculator dummy is equal to 1 if an investor displays an above-median level of APROP on average during the pre-period and zero otherwise. Year zero refers to the year during which we first classify an investor as being socially connected. Two investors are defined as being connected during a calendar year if, on at least 5 days during the year, they purchased the same stock. The sample period is 1997 to 2011. T-statistics are shown in parentheses and ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	Proportion in highest VOL quintile		**Proportion in highest IVOL quintile**		**Proportion in highest SKEW quintile**	
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-connection average network APROP	0.28771***	0.0939***	0.2538***	0.1396***	0.2548***	0.0368
	[10.84]	[2.60]	[11.26]	[4.54]	[2.84]	[0.78]
Higher risk investor dummy		0.0514***		0.0319***		0.0323***
		[8.34]		[6.85]		[6.31]
Pre-connection average network APROP x Higher risk investor dummy		0.3412***		0.2097***		0.3410***
		[6.54]		[4.70]		[5.50]
Female	0.0100	0.0097***	0.0056	0.0047	0.0087	0.0097
	[1.16]	[1.16]	[0.88]	[0.75]	[1.22]	[1.38]
Age	-0.0001	0.0001	-0.0000	0.0001	0.0004*	0.0005**
	[-0.45]	[0.59]	[-0.25]	[0.76]	[1.86]	[2.22]
Log(Average trade value)	-0.0130***	-0.0130***	-0.0140***	-0.01340***	-0.0120***	-0.0114***
	[-4.84]	[-4.59]	[-7.05]	[-6.84]	[-5.39]	[-5.20]
Trade frequency	0.0004***	0.0004***	0.0002**	0.0002**	0.0002*	0.0002**
	[3.31]	[3.10]	[2.19]	[1.97]	[1.89]	[2.01]
N	2,034	2,034	2,034	2,034	2,034	2,034
R-squared	0.0795	0.1223	0.0918	0.1181	0.0527	0.0787

Table 5: Regression – Moderating effects of social interaction for conservative and speculative investors (Changes)

This table displays coefficient estimates for OLS regressions where the dependent variable is the change in an investor's APROP between the average across the years prior to event year zero and year zero. APROP refers to the abnormal proportion of an investor's total trade value accounted for by stocks in the highest risk quintile during event year zero year where risk is measured by volatility, idiosyncratic volatility and skewness. Pre-connection difference in APROP refers to the difference between the average values of APROP during the pre-connection period for the individual and for the individuals comprising the investor's social network. The speculator dummy is equal to 1 if an investor displays an above-median level of APROP on average during the pre-period and zero otherwise. Year zero refers to the year during which we first classify an investor as being socially connected. Two investors are defined as being connected during a calendar year if, on at least 5 days during the year, they purchased the same stock. The sample period is 1997 to 2011. T-statistics are shown in parentheses and ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	Change in proportion in highest VOL quintile		**Change in proportion in highest IVOL quintile**		**Change in proportion in highest SKEW quintile**	
	(1)	(2)	(3)	(4)	(5)	(6)
Pre-connection difference in APROP	-0.2764***	-0.0911***	-0.3833***	-0.1747***	-0.3006***	-0.1070***
	[-13.49]	[-2.95]	[-19.57]	[-5.30]	[-12.91]	[-2.70]
Speculative investor dummy		-0.0860***		-0.0573***		-0.0583***
		[-9.04]		[-7.29]		[-8.26]
Pre-connection difference in APROP x Speculative investor dummy		-0.2013***		-0.2265***		-0.1763***
		[-4.64]		[-5.21]		[-3.47]
Log(Average trade value)	0.0062	0.0028	0.0005	-0.0019	-0.0026	-0.0049*
	[1.64]	[0.75]	[0.15]	[-0.66]	[-0.93]	[-1.79]
Trade frequency	0.0006***	0.0006***	0.0003*	0.0003*	0.0005***	0.0005***
	[2.97]	[2.97]	[1.70]	[1.67]	[2.91]	[3.04]
N	1,364	1,364	1,364	1,364	1,364	1,364
Adjsuted R-squared	0.1262	0.1904	0.2216	0.2689	0.1136	0.1658

Table 6: Regression - Relationship between investment in high-risk stocks and social interaction

This table displays coefficient estimates for the following regression equation.

$$APROP_{i,t} = \propto + \beta_1 POST_t + \beta_3 TRADING\_FREQ_{i,t} + \text{INVESTOR FIXED EFFECTS} + \varepsilon_{i,t}$$

$APROP_{i,t}$ refers to the abnormal proportion of an investor's total trade value accounted for by stocks in the highest risk quintile during a given year where risk is measured by volatility, idiosyncratic volatility and skewness. $POST_t$ is a dummy variable set equal to 1 from the year of first connection onwards and zero otherwise. $TRADING\_FREQ_{i,t}$ is the number of stock-trade days during the year on which an investor was active. Two investors are defined as being connected during a calendar year if, on at least 5 days during the year, they purchased the same stock. The sample period is 1997 to 2011. T-statistics are shown in parentheses and ***, ** and * denote statistical significance at the 1%, 5% and 10% levels respectively.

	DV = APROP measured by volatility	**DV = APROP measured by idiosyncratic volatility**	**DV = APROP measured by skewness**
	(1)	(2)	(3)
Post-connection dummy	-0.0008	-0.0071**	-0.0080***
	[-0.25]	[-2.43]	[-2.86]
Trading frequency	-0.0001	0.0005	0.0000
	[-0.78]	[0.58]	[0.40]
Investor fixed effects	Yes	Yes	Yes
N	13,440	13,440	13,440
R-squared	0.3336	0.3365	0.2459

Table 7: Regression – Calendar time portfolio regression alphas

This table displays coefficient estimates from a calendar time portfolio regression estimated as follows. We divide investors into three groups based on their quintile ranking of APROP during the pre-connection period. Group 1 contains investors falling into quintile 1, group 2 contains those falling into quintiles 2-4 and group 3 contains those falling into quintile 4. On each trading day, we further divide the investors contained in each risk group into a connected and a non-connected subgroup. Investors move into the connected subgroup on the first day of the year in which they are first identified as connected. For each group and subgroup of investors, we form a portfolio based on the investors' prior stock purchase decisions. On each trading day, the portfolio takes a long position in the stocks for which the investors in the group who were net buyers of the stock during the last 90 calendar days outnumber those who were net sellers. We then regress the daily return of this portfolio against the three Fama French factors to estimate the risk-adjusted daily alpha. $\mathrm{APROP}_{i,t}$ refers to the abnormal proportion of an investor's total trade value accounted for by stocks in the highest risk quintile during a given year where risk is measured by volatility, idiosyncratic volatility and skewness. The sample period is 1997 to 2011. T-statistics are shown in parentheses.

Panel A: APROP measured by volatility

		1-day α (%)	RM - RF	SMB	HML	N	R-squared
Risk group 1:	Unconnected	0.0290	0.7901	0.1641	0.0162	3578	0.7675
		[3.08]	[99.89]	[13.32]	[1.39]		
	Connected	0.0278	0.9106	0.2981	-0.1788	3028	0.7508
		[2.16]	[84.76]	[17.43]	[-10.88]		
Risk group 2:	Unconnected	0.0214	0.8239	0.3327	-0.0599	3579	0.7570
		[2.18]	[99.45]	[25.78]	[-4.91]		
	Connected	0.0245	0.9122	0.3659	-0.2470	3034	0.7730
		[2.01]	[90.82]	[23.15]	[-16.23]		
Risk group 3:	Unconnected	0.0168	0.9267	0.4397	-0.1476	3577	0.7323
		[1.42]	[92.51]	[28.19]	[-10.02]		
	Connected	0.0253	1.0789	0.5390	-0.4208	3028	0.6751
		[1.35]	[69.26]	[21.74]	[-17.65]		

Panel B: APROP measured by idiosyncratic volatility

		1-day α (%)	RM - RF	SMB	HML	N	R-squared
Risk group 1:	Unconnected	0.0268	0.8088	0.1648	0.0065	3578	0.7678
		[2.78]	[99.70]	[13.04]	[0.54]		
	Connected	0.0209	0.8864	0.2560	-0.1573	3032	0.7642
		[1.72]	[87.58]	[15.91]	[-10.17]		
Risk group 2:	Unconnected	0.0206	0.8133	0.2945	-0.0604	3582	0.7596
		[2.12]	[99.49]	[23.13]	[-5.02]		
	Connected	0.0223	0.9467	0.4285	-0.3160	3031	0.7475
		[1.63]	[84.09]	[24.18]	[-18.52]		
Risk group 3:	Unconnected	0.0183	0.9308	0.4741	-0.0966	3576	0.7207
		[1.51]	[91.24]	[29.83]	[-6.43]		
	Connected	0.0529	1.0337	0.4790	-0.3813	3034	0.6634
		[2.88]	[68.00]	[20.24]	[-16.74]		

Panel C: APROP measured by skewness

		1-day α (%)	RM - RF	SMB	HML	N	R-squared
Risk group 1:	Unconnected	0.0157	0.8503	0.1818	-0.0592	3578	0.7550
		[1.48]	[95.33]	[13.08]	[-4.51]		
	Connected	0.0200	0.9885	0.3121	-0.2228	3031	0.7406
		[1.38]	[81.85]	[16.24]	[-12.06]		
Risk group 2:	Unconnected	0.0238	0.8043	0.3104	-0.0561	3582	0.7654
		[2.53]	[101.57]	[25.17]	[-4.81]		
	Connected	0.0290	0.8908	0.3091	-0.2647	3032	0.7543
		[2.28]	[84.72]	[18.89]	[-16.79]		
Risk group 3:	Unconnected	0.0281	0.8673	0.4166	-0.0293	3577	0.7387
		[2.62]	[96.34]	[29.71]	[-2.21]		
	Connected	0.0362	0.9596	0.4460	-0.2224	3032	0.7083
		[2.44]	[78.18]	[23.09]	[-11.96]		

Figure 1: Percentage of connected buy trades by trading day

We identify all the connected buy trades for each stock: instances where two investors purchase the same stock on the same day. For each stock which has at least 100 connected buy trades during the year, we then compute the percentage of the total yearly connected buy trades that occur on each trading day. We then rank the trading days according to this percentage value within each stock-year and compute the average for each ranked trading day. The chart plots these averages for ranked trading days 1 through 50. The sample period is 1997 through 2011.

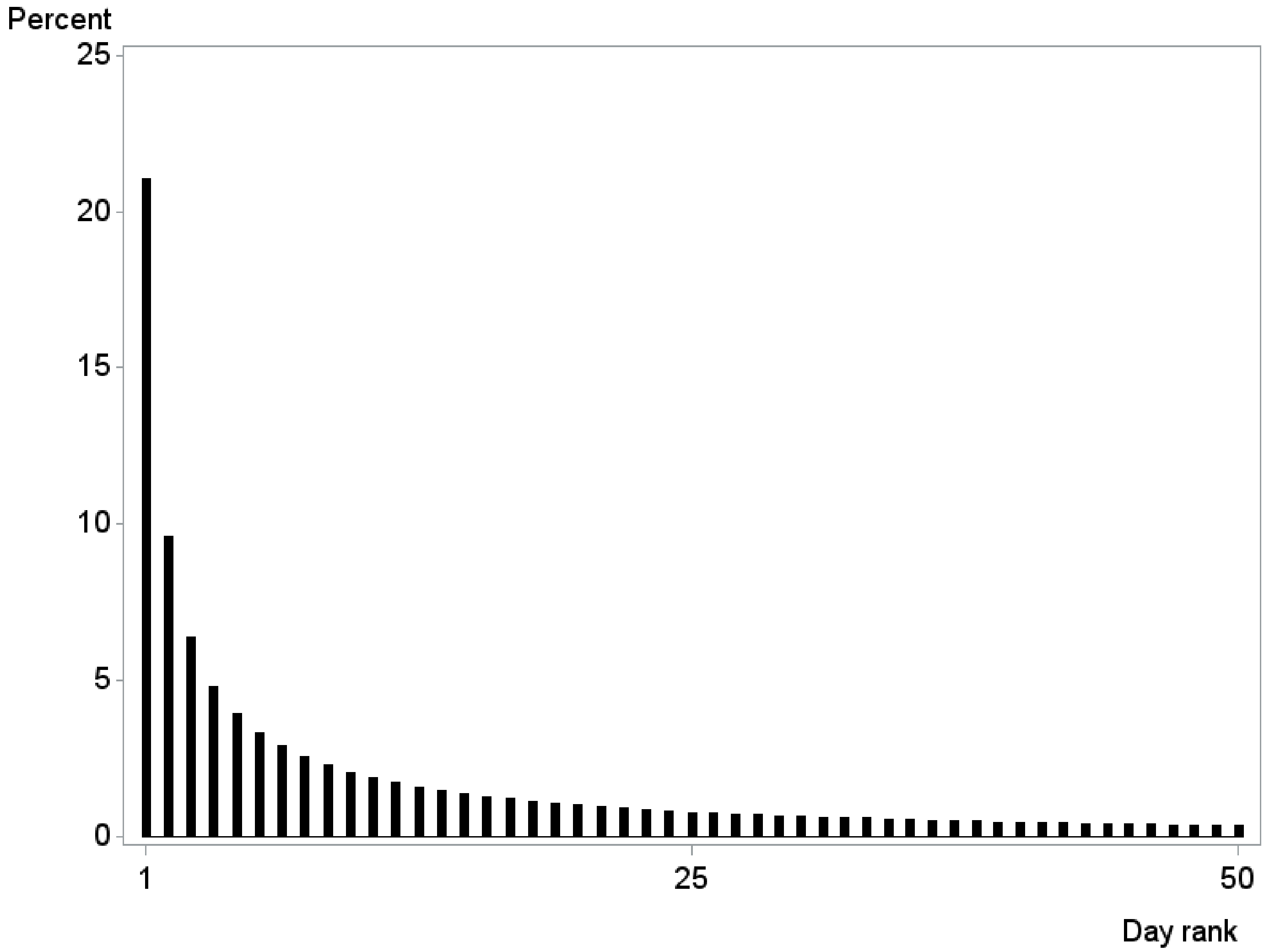


Figure 2: Expected and actual CDFs of distances between residences of socially connected investors

The black broken line depicts the expected cumulative distribution function (CDF) of distances (in km) between residences of socially connected investors during 2011. The expected CDF is computed assuming that social network connections between pairs of investors occur at random within the entire sample of connected investors. The red solid line depicts the actual CDF of distances between residences of socially connected investors. Two investors are defined as being connected during a calendar year if, on at least 5 days during the year, they purchased the same stock.

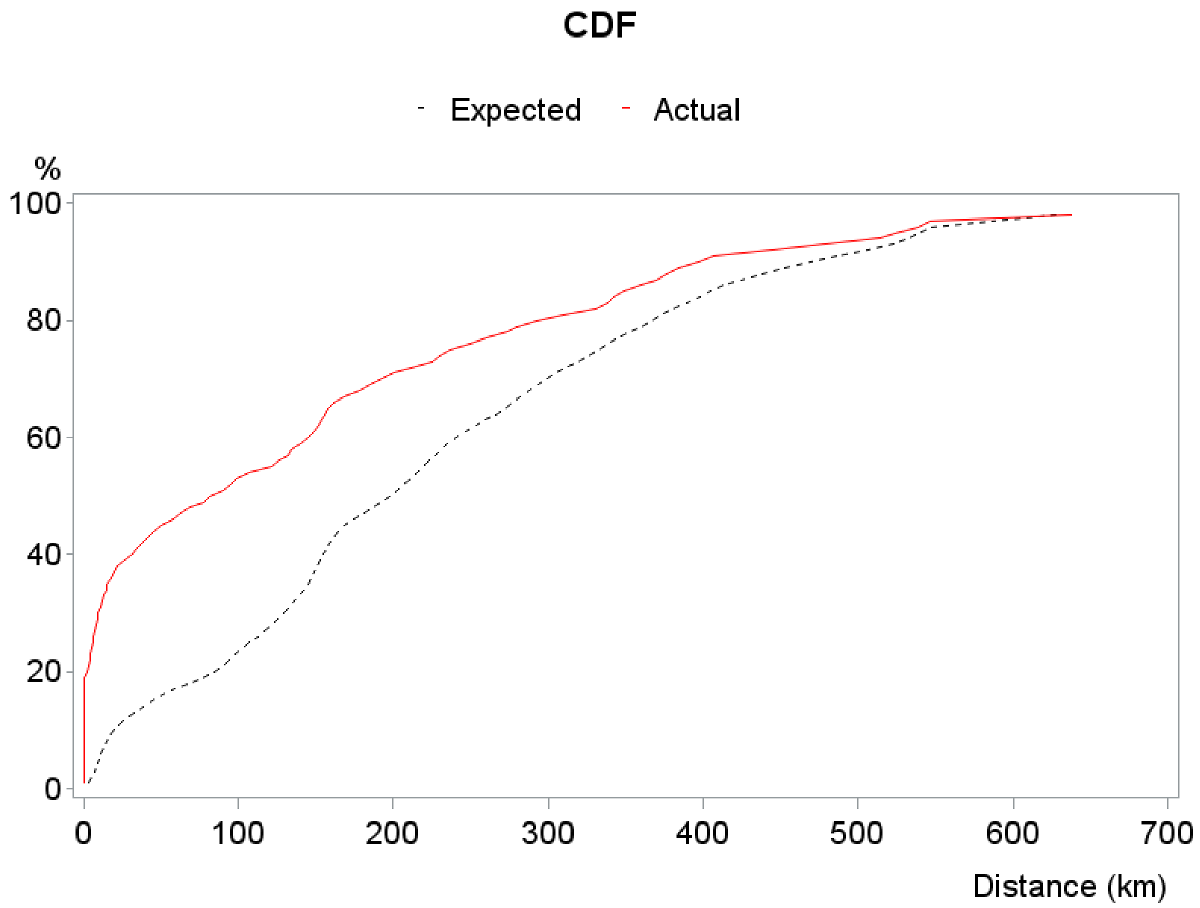
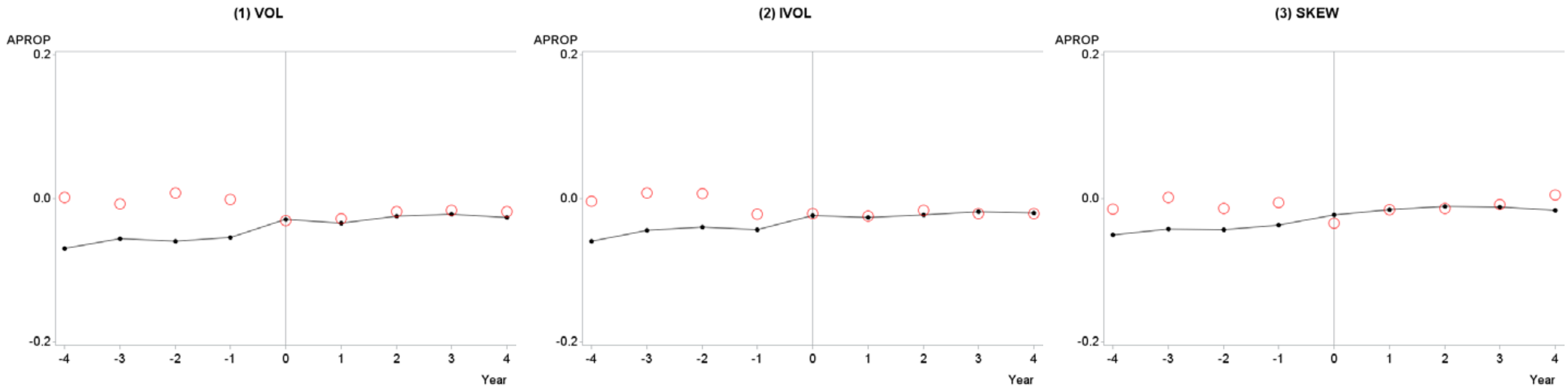


Figure 3: Moderating effects of social interaction for conservative and speculative investors

The black solid lines depict levels of APROP for investors active during each event year between year -4 and year +4. The red circles depict levels of APROP averaged across the individuals in a given investor's social network (identified during year zero). APROP refers to the abnormal proportion of an investor's total trade value accounted for by stocks in the highest risk quintile during a given year where risk is measured by volatility, idiosyncratic volatility and skewness (specifications 1, 2 and 3 respectively). Event year zero refers to the year during which we first classify an investor as being socially connected. Two investors are defined as being connected during a calendar year if, on at least 5 days during the year, they purchased the same stock. Results are shown separately (in Panels A and B respectively) for investors who display a below- (conservative) and above-median (speculative) average value of APROP during the pre-connection period. The sample period is 1997 to 2011.

Panel A: Conservative investors



Panel B: Speculative individuals