working paper

15 22

# Facebook Finance: How Social Interaction Propagates Active Investing

Rawley Z. Heimer and David Simon

FEDERAL RESERVE BANK OF CLEVELAND

**Working Paper 15-22** October 2015

**Facebook Finance:**
**How Social Interaction Propagates Active Investing**

Rawley Z. Heimer and David Simon

This paper shows how active investing strategies propagate through social connections in a network of retail traders, using a new database of social activity linked to individual-level trading records. A trader's good short-term performance causes them to contact others. A trader's activity increases when peers perform well and increase communication. We use the staggered entry of brokerages into partnerships with the social networking platform, which is a necessary precursor for traders to access the network, to argue these effects are causal. This pattern of communication supports active trading, even though the network reveals the low success rate of retail traders.

Suggested citation: Heimer, Rawley Z., and David Simon, 2015. "Facebook Finance: How Social Interaction Propagates Active Investing," Federal Reserve Bank of Cleveland, working paper no. 15-22.

Rawley Z. Heimer is at the Federal Reserve Bank of Cleveland and David Simon is at Berkeley Research Group. Heimer can be contacted at: rawley.heimer@researchfed.org.

# 1 Introduction

Some individuals trade actively to try and beat the market. For instance, male traders who use a discount brokerage turn over their portfolio every 1.3 years on average (Barber and Odean (2001)). Despite their efforts, active traders have historically underperformed relative to passive benchmarks. An investor who switched to a passive market portfolio between 1980 and 2006 could have averaged an additional 67 basis points per year (French (2008)). The high trading volume of active investors is difficult to reconcile with a canonical model of investing under uncertainty. Hence, this behavior seems puzzling.

This paper argues that social interaction helps explain the popularity of active strategies. We provide empirical support for recent theoretical research (Han and Hirshleifer (2013)) motivated by the potential ways in which people communicate about their investments and respond to news about investment returns. In particular, individuals prefer to broadcast their investment successes while remaining silent after losses. The better the performance of traders who talk, the more likely the recipients of communication are to adopt similar trading strategies. Owing to the increase in portfolio variance inherent to active strategies, those who trade actively have more opportunity to selectively discuss extreme positive returns and are therefore more persuasive. Recipients of communication are then more likely to misguidedly trade actively though not necessarily more profitably.

To study how social interaction relates to active trading, we introduce new data from a sample of retail foreign exchange traders who are members of a Facebook-style social network that for the purpose of anonymity, we call myForexBook.[1] Prior to joining, traders must have

[1]The retail foreign exchange market, which did not exist even a decade ago, has grown tremendously since the advent of online trading. According to King and Rime (2010), worldwide retail foreign exchange trading volume grew over 70% from 2007 to 2010 and now exceeds $125 billion per day, roughly the same as daily turnover on the entire NYSE family of stock exchanges (NYSE, Arca, and Amex).

an account on one of roughly 50 online brokerages from which myForexBook collects trading activity in real-time. The database contains the detailed trading history and communications of several thousand traders collected between early 2009 and December 2010. It includes over two million time-stamped trades and over one hundred thousand time-stamped messages and friendships, which allow us to identify clear links between trading and social activity. Also, our data avoids concerns about reporting bias because the social-networking website extracts trading records directly from a trader's brokerage account in real time, a feature which offers an advantage over studies that collect postings from internet forums and chatrooms (Antweiler and Frank (2004), among others). Despite the uniqueness of the empirical setting, we present a favorable comparison to individual investors studied more widely in the finance literature, namely the discount brokerage data used in Barber and Odean (2000) and the population of Finnish stock traders (Grinblatt and Keloharju (2000)).

The patterns of communication within the network can be illustrated by the example of trader 478, a 34-year-old Indonesian who made 792 trades over the course of ten weeks. One week in November 2009, trader 478 made a total profit of $72,303, becoming one of the most successful traders that month. Trader 478 sent messages to 128 other traders that week, more than twice the number contacted during the previous week, presumably to celebrate his or her performance. Upon receiving these messages, the 128 recipients increased their aggregate trading activity the following week by almost 30%, from 1,707 trades to 2,148 trades, without experiencing any appreciable gains in profitability.

These empirical patterns can be interpreted causally. New traders enter myForexBook gradually over the sample period, which enables us to use a panel-data analysis to compare trader activity before and after exposure to the social network. Moreover, a trader is unable to join the network until his or her brokerage has reached legal and technological agreements with myForexBook. The staggered incorporation of new brokerages is similar to an instrumental variable that predicts trader entry, but that empirical evidence suggests is likely exogenous to the behavior of any individual or group of traders. This allows us to isolate the causal influence of social interaction from other contemporaneous factors, thereby alleviating well-known concerns over endogeneity and reflection that have made it difficult for researchers to use social network data to develop interpretable empirical results (Manski (2003)).

A rigorous set of difference-in-difference tests utilize this insight about trader entry to show that individuals trade 20% more when their peers have positive returns. The effect is stronger when traders receive messages from others and at times when there is more conversation in the network. On the other hand, poor peer performance does not deter trading, a surprising result, because losses are informative about the average success rate of retail traders. Also, a placebo test using false dates of network-entry provides evidence that secular trends formed prior to joining myForexBook are unlikely to explain these findings. Furthermore, communication within the network is biased towards positive returns. An increase from the 10th to the 90th percentile in the value of a trader's portfolio causes about a five percentage point increase in the probability that he or she will contact other traders.

To analyze the implications of these results, we present suggestive evidence that the social network contributes to aggregate trends toward active trading. Average trading intensity and the variance of portfolio returns have both increased since the inception of the online network. The channels of communication within the network imply that the most active traders encourage others to increase their activity, while contemporaneous market factors are unlikely to explain these findings. Taken as a whole, our analysis supports the conclusion that social interaction propagates active investing.

This paper makes an important contribution to the empirical literature on social interaction in finance by being among the first to go beyond documenting the influence of social interaction to exploring the mechanisms by which investment ideas spread. Furthermore, our study is the first to our knowledge to examine instances of observed communication, rather than relying on shared geography or background to infer social interaction. Despite the possibility that proxies typically used in the literature reflect other individual-level similarities that could potentially influence subsequent outcomes, empirical studies have attempted to use social interaction to explain a variety of phenomena in the finance literature. It helps alter savings decisions (Duflo and Saez (2003) and Beshears et al. (2011)), promotes stock market participation (Hong, Kubik, and Stein (2004), Brown et al. (2008), and Li (2014)), and potentially influences investment choices (Ivković and Weisbenner (2007) and Kaustia and Knüpfer (2012)).[2] Highly connected traders can explain the heterogeneity in trader performance (Bildik et al. (2014)). Several papers present related findings among mutual fund managers (Hong, Kubik, and Stein (2005), Cohen, Frazzini, and Malloy (2008), and Pool, Stoffman, and Yonker (2014)). Social interaction even affects firm policies and governance (Fracassi and Tate (2012), Shue (2013), and Popadak (2014)) and relationships with banks (Engelberg, Gao, and Parsons (2012)).

It is important to understand what drives the participation of active investors in financial markets since they can have a profound effect on market outcomes as well as their own welfare. It has been widely documented since Barber and Odean (2000) that active retail investors lose relative to passive benchmarks. More recently, Barber et al. (2009) find that Taiwan's retail investors underperform compared to the market by 3.8% and accumulate losses that amount to 2.2% of Taiwan's GDP annually. To explain these findings, the litera-

[2] Our findings are likely most comparable to Kaustia and Knüpfer (2012), who find a similar discontinuity between geographically localized returns and stock market entry.

ture often points to investor overconfidence (Barber and Odean (2001a), Bénabou and Tirole (2002), Biais et al. (2005), and Grinblatt and Keloharju (2009)). Individual investors potentially exhibit nonstandard preferences (Kumar (2009)). A few recent explanations suggest that the observed underperformance is evidence of rational traders learning about their own skill (Linnainmaa (2011)) or the risk of adverse selection (Linnainmaa (2010)). Our results do not preclude these other theories.

The paper is organized as follows. Section 1 describes the theoretical foundation for this research. Section 2 presents the social networking data, and Section 3 describes our identifying assumptions. Section 4 shows that receivers of communication increase their activity in response to hearing of high returns, while Section 5 shows that short-term returns lead to increased communication. Section 6 presents the implications of these empirical findings by demonstrating that the social network has helped propagate active investing strategies. Section 7 concludes.

## 2 Theory: Social networks in finance

A wealth of theoretical models examine how heterogeneous information provision within networks of market participants affect equilibrium asset market outcomes (Grossman (1976) and Hellwig (1980), among others). A common theme emerges: if diverse information is interpreted rationally and information-flow is unbiased, prices communicate all necessary information. Therefore, trading is motivated by shocks to fundamentals (or other parameters), leaving little room for communication within social networks or the network topology to matter.

In contrast, *Self-Enhancing Transmission Bias and Active Investing* (Han and Hirshleifer (2013)) shows that communication biases play a role in investment decision-making and potentially contributes to our understanding of the active-investing puzzle. To motivate our empirical tests, we highlight a few features of the Han and Hirshleifer (2013) model.

In Han and Hirshleifer's (2013) model, there is a population of $n$ traders, $i$, who take on one of two investment styles, Active or Passive, denoted $A$ and $P$ respectively. Active traders pursue hands-on strategies that require more trading, which includes opportunity or transaction costs, $D$. Meanwhile, Passive traders can be thought of as costlessly holding the market portfolio. Active strategies are characterized as having higher variance, $\sigma_i^2$, and are more sensitive to common factors, $\beta_i$, but do not necessarily have higher returns, $R_i$.

When two randomly drawn traders meet, the Sender chooses whether or not to discuss his or her returns, and does so with probability $s(R_i)$. **Assumption 1:** Traders exhibit self-enhancing transmission bias, the tendency to broadcast successes while downplaying failures. This implies that the probability a Sender reveals his or her strategy's performance is increasing in his or her returns,

$$\textbf{Sender Function}: \quad s(R_i) = \lambda R_i + \gamma \tag{1}$$

with $\lambda > 0$. The baseline probability of transmission is $\gamma \in [0, 1]$, which reflects average investor sociability.

If and when another trader, termed the Receiver, learns of the Sender's returns, they exhibit a probability, $r(R_i)$, of adopting the Sender's strategy. Han and Hirshleifer (2013) call this the Receiver function:

$$\textbf{Receiver Function}: \quad r(R_i) = aR_i^2 + bR_i + c \tag{2}$$

**Assumption 2:** Recipients learn about strategies through their communications and therefore believe that the performance of the strategy discussed by the Sender reflects the true distribution of $R_i$. Thus, Receivers do not adjust for the bias in Sender communications, and the probability of the Receiver's conversion is increasing in the Sender's returns, $b > 0$. Han and Hirshleifer (2013) include a quadratic term with a positive coefficient, $a > 0$, to allow extreme returns to be more persuasive, although most of the model's implications can be obtained without this feature of the Receiving function. The baseline conversion rate is also positive, $c \in [0, 1]$.

Han and Hirshleifer (2013) shows that the expected change in the fraction of Active traders, $f_A$, is increasing in the baseline sociability of traders (see Appendix A.1 for the algebra),

$$\left(\frac{2n}{\chi}\right) \frac{\partial E\left[\triangle f_A\right]}{\partial \gamma} = a\left((\beta_A^2 - \beta_P^2)\sigma_r^2 + (\sigma_A^2 - \sigma_P^2)\right) - bD + aD^2 > 0 \tag{3}$$

so long as the costs associated with active trading are sufficiently small ($\chi$ is the probability that an $A$ and $P$ pairing is drawn at random).[3]

We provide causal empirical tests of the assumptions behind the Sending and Receiving functions. The two assumptions combine to imply that increased sociability leads to more active trading, the comparative static presented in Eq. 3. The creation of online social networks which enable investors to communicate with one another can be considered a shock to aggregate trader sociability, $\gamma$.

[3]Retail forex traders pay the spread on all spot transactions. Since the spread is the only trading cost and they tend to average only a few pips, it seems reasonable to presume that the parameter $D$ is small in our setting.

## 3 The myForexBook network: Description and data

The data used in this research was compiled by a social networking website that, for the purposes of anonymity, we call myForexBook. Registering with myForexBook requires a trader to have an open account with one of 53 retail specific foreign exchange brokers with which myForexBook maintains a technological and legal agreement.[4] Upon joining the social network, a trader sets up his or her user profile, an example of which is displayed in Fig. 1. Traders cannot use myForexBook's web-platfrom to initiate or close trades, but new positions entered on a trader's brokerage account are recorded by myForexBook and are time-stamped to the second. Hence, our data has an advantage over existing studies of stock-market-message-board postings, because there are no concerns about reporting bias (Antweiler and Frank (2004), Chen et al. (2014), and Mizrach and Weerts (2009)).[5]

The myForexBook web platform offers an unprecedented opportunity for traders to communicate about their investments through features such as a discussion forum and a peer-to-peer messaging platform. Also, after forming a bilateral friendship that is initiated via a friend request, registered users are able to view each others' trading activity in real time, a feature illustrated in Fig. 2. Since myForexBook extracts trading activity directly from its

[4]Professional traders housed at banks or other institutions place trades via routing services which can accommodate large positions such as Electronic Broking Services (EBS) or Reuters 3000 Xtra (King and Rime (2010)). Professional traders and the routing services do not have any affiliation with myForexBook.

[5]Studies of traders that are among the first to utilize online platforms are likely comparable to this research ((Barber and Odean (2001b) and Choi, Laibson, and Metrick (2002)).

members' brokerage accounts, communication between traders is unlikely to be contaminated by "cheap talk."[6,7]

Our slice of myForexBook data begins in January 2009 and extends to early December 2010, with the social networking features of the data starting around May 2009. The database includes daily account balances per user and, after cleaning, 2,149,083 opened positions, of which 2,144,357 had been closed. We then restrict our data to traders who issue at least 25 trades and who we observe at least ten weeks of trading activity. This produces an unbalanced panel of 3,117 traders over a total of 111,928 weeks. Our results are robust to alternative data trimmings, but lowering the threshold for inclusion contributes little to the panel-data analysis used herein.

### *3.1. Portfolio returns*

We use our data to generate weekly returns per trader $i$. Considering that a substantial amount of the activity in the forex market centers around the release of economic news and less than two percent of all trades are issued on weekends, week-to-week returns best characterize the activity of these traders. Furthermore, other empirical studies of information provision in financial markets emphasize the use of weekly as opposed to daily or monthly returns to minimize noise while retaining inference (Hou and Moskowitz (2005)). However, our results are robust to a daily analysis (available upon request).

[6] Note that several of the online trader social networks have recently begun to offer a feature that allows traders to automatically copy the trades of their friends. There are no such trades in our data set, because the technology was not introduced on the myForexBook website until several months after our sample ends. Furthermore, to the best of our knowledge, this technology was not available anywhere on the world wide web until well after the sample period in our data.

[7] myForexBook did not provide us with the content of the user messages, just the time and the direction of the communication.

We follow Barber and Odean (2000) and calculate the weekly gross return on the trader's portfolio,

$$R_{i,t} = \sum_{k=1}^{s_{i,t}} p_{k,t} \cdot R_{k,t} \tag{4}$$

where $p_{k,t}$ is the value of position $k$ when it is opened at second $t$ divided by the total opening value of all positions held by trader $i$. $R_{k,t}$ is the return on position $k$ and $s_{i,t}$ is the number of positions opened by $i$. Although the calculation includes transaction costs, they are unlikely to have much impact on $R_{k,t}$, because retail brokerages in the forex market charge only half of the bid-ask spread per transaction and spreads tend to average no more than one or two pips per trade. The empirical results are robust to using short-term government bonds to calculate excess returns, $R_{i,t} - R_t^f$, but this has a negligible effect on the returns calculation because risk-free rates were at historic lows during the sample period (available upon request). Moreover, the regression analysis includes weekly fixed effects, which captures variation in aggregate market conditions and performs a similar role to $R^f$. We also Winsorize $R_{i,t}$ at the outer one percent of the distribution.

Traders in the myForexBook database lose 0.028 per week with a standard deviation of 0.20 (Table 1, Panel I). These results are comparable to existing studies of retail traders in equities (Barber and Odean (2000)). Retail investors at a discount brokerage lose roughly 0.05 per month in their common stock investments between 1991 and 1996. While the traders in our sample appear to lose even more if losses are compounded monthly, much of the difference between our sample and the retail stock traders in Barber and Odean (2000) can be attributed to the greater availability of leverage in the retail forex market during the sample period. Traders in our sample average 8.6 times leverage per trade, which if reduced to the levels in regulated equity markets, would place the returns in our sample on par with comparable studies. Furthermore, our sample is similar to the U.S. population of retail forex traders. According to quarterly reports compiled by the Commodity Futures Trading Commission (CFTC), roughly a third of all brokerage accounts are profitable.

### *3.2. Trading volume*

It is challenging to find comparisons to the trading volume in our sample, because this is among the first studies of retail foreign exchange traders. However, there is evidence that retail trading is sensitive to trading costs (Foucault, Sraer, and Thesmar (2011) and Ben-David, Heimer and Hou (2014)), which tend to be low in the forex market. Therefore, the institutional features of the forex market – no fixed transaction costs and highly liquid prices – should be expected to produce higher baseline rates of turnover than among comparable participants in equity markets.

Our outcome variable of interest, $trade.count_{i,t}$, is the number of trades executed by trader $i$ per week $t$. (Table 1, Panel I). Traders average around 25 trades per week with a median of 11. Our estimation results are robust to weighting each trade by its value divided by the value of the portfolio (available upon request). We have also explored the use of leverage as an outcome variable. However, regulatory changes imposed by the CFTC during our sample period, which were targeted at leveraged trading in the U.S., make it difficult to interpret the results.

For the convenient purposes of this research, the total value of trading volume originating from the social network is diminutive in comparison to the size of the aggregate forex market. This is not to suggest this study is not meaningful, because according to estimates from the Bank of International Settlements (King and Rime (2010)), the aggregate market averages \$4 trillion in daily volume, with the retail share constituting around ten percent. Meanwhile, the volume of trading in the entire two years of myForexBook data, is approximately $125 billion to $150 billion, or less than half of one day's worth of trading by the entirety of the retail market. Therefore, it seems unlikely that the social network can be used to manipulate or influence prices. This alleviates concerns that forex price changes are endogenous to the network, which in turn suggests that reverse causality is an unlikely alternative explanation. It enhances our confidence in using idiosyncratic portfolio shocks – the deviation in returns after accounting for unobservable ability via trader fixed effects – as an exogenous variable.[8]

### *3.3. Social interaction*

Our data contains the complete record of social activities within myForexBook, including the times of logins and messages sent, and friendships established and rejected. Many of these statistics are without known benchmarks. However, evidence that the social inclinations of myForexBook traders are comparable to that of retail investors in more widely used data sets builds confidence in the representativeness of our findings. Appendix A.2 replicates and finds results that are similar to the epidemic model of Shive (2010), in which the probability of trading depends on the fraction of stock market participants in Finland who have previously traded the asset.

Throughout the analysis, we employ different configurations of a trader's peer group $j$ to exploit the dynamics of the network-building process and the possibility that there is variation in the strength of connections. However, since friends are able to view each other's trading, our most common definition is that a trader belongs to $j$ if they develop a friendship with $i$ at any point by the end of the sample. To form a bilateral friendship, one trader submits a friend request and the other trader accepts, although some requests are

[8]Engelberg and Parsons (2014) offer a comparable example of stock market returns causing personal behavioral changes.

denied. Trader friendships can only form after both traders have joined myForexBook, and the data contains the date at which each friendship is reached. The average trader has 21.8 friends ($peer.group.size_i$) by the end of the sample (Table 1, Panel II).[9]

Figure 3 presents the complete set of connections in the network. The number of friends per trader is represented by the size of each node. The image uses a visualization algorithm to show that certain groups of traders – represented by a color spectrum – are more tightly connected to each other than they are to others (Blondel et al. (2008)). Clearly, a distinct network structure has evolved since myForexBook's inception.

We utilize the personal messaging feature of myForexBook to study revealed attempts to contact other traders. Traders are able to send peer-to-peer messages to one another ("Send Private Message" in the upper right corner of Fig. 1). The average trader has sent (received) 28.2 messages during his or her trading career. The distribution of received messages (median equal to 19.0) is more uniform than the distribution of sent messages (median equal to 17.0). These social networking statistics appear to follow a power-law distribution, which is a common feature of social networking studies in nonfinancial settings.

### *3.4. Additional control variables*

The database also contains information on the characteristics of its members, collected when traders create an account on myForexBook (Table 1, Panel III). The median trader in our database is 36.2 years old and lives in either the USA or Western Europe. They have one to three years of trading experience and classify their trading approach as technical as opposed to fundamental, news-based, or momentum-based.[10] This information is constant for each

[9]We do not include social interaction with the set of traders excluded from our sample.

[10]To examine response accuracy, the retail brokerages in the sample provide the location of around 70% of all accounts. The trader survey responses regarding location have close to a perfect match to the locations provided by the brokerage.

trader and is thus collinear to trader fixed effects, which are included in most of the analysis. However, the empirical results are robust to the inclusion of these variables as controls in OLS regressions estimated without fixed effects (available upon request).

We control for common factors causing variation in aggregate returns specific to the forex market and to the social network. The first benchmark, $DXY_t$, provided by Bloomberg, is a daily index of the spot value of the U.S. dollar relative to a weighted basket of nine other currencies, from which we calculate the percentage change between consecutive Wednesdays. The second is $CVIX_t$, an implied volatility index for the currency markets produced by Deutsche Bank. Aggregate factors could also drive aggregate conversation in the social network. Thus, $average.chatter_t$ is an indicator variable equal to one if the the average number of messages per user sent within the network during week $t$ is greater than the moving average of five weeks before and after.

## 4 Identification: Agreements between brokerages and myForexBook

For a trader to join myForexBook, his or her brokerage must have first partnered with myForexBook. As illustrated in Fig. 4, new brokerages partnered with myForexBook gradually over the course of the sample period. This staggered process was driven by legal and technological agreements between myForexBook's operators and partnering brokerages. To offer more detail, myForexBook extracts confidential trading records in real time from a selection of brokerages, all of which have a unique database infrastructure. This means that myForexBook is not only required to reach a nondisclosure agreement with the brokerage, but it also has to make its software compatible with the structure of the brokerage's server.[11]

[11]Providing a discrete example that includes the names of one or more retail brokerages would potentially compromise the identity of our data provider.

The data contains an indicator variable for the brokerage used by the trader and the date at which each new trader enters the network. The former enables the use of brokerage fixed-effects to account for any brokerage-specific factors that could confound the relation between trader entry and trading behavior. The latter allows us to update, when necessary, the set of traders that belong to $i$'s peer group $j$ as it evolves over the sample period, with confidence that the timing of additions is quasi-random. This is important because random assignment to peer groups counteracts the reflection problem and other endogeneity concerns within social networks (Manski (1993)).

The database also contains trading records from prior to entering the network, which constitutes 41% of all trades. This feature enables a comparison of trading activity before and after $i$ gains access to the social networking features of the web platform. For example, to estimate the Receiver's function, a difference-in-differences framework compares traders who join the social network to traders who are similar, but who are excluded from joining myForexBook at any given point in time. This helps isolate the causal influence of social interaction from contemporaneous factors that potentially confound inference.

Empirical evidence offers support for our identifying assumptions. The incorporation of new brokerages is a strong predictor of the time at which a trader joins myForexBook. An OLS regression of a trader's join date on the set of brokerage dummy variables produces an F-statistic of 352.

Additionally, the process by which new brokerages were added to myForexBook is likely uncorrelated with the characteristics of any individual or group of traders. This is important for identification because traders that are the first to join myForexBook can generally be thought of as being part of a treatment group, while traders who are excluded from joining myForexBook until late in the sample are more often part of the control group. Appendix

A.3 and Table 2 (Panel I) provide a comparison of early and late entrants into myForexbook, and find that the two groups are not statistically different. Also in Table 2 (Panel II), several Probit models provide evidence that observable trader characteristics cannot explain which traders are the first to join myForexBook. Taken together, this suggests that brokerage agreements are free of selection bias and other confounds that would invalidate its use as an unbiased predictor of exposure to the social network.

# 5 The Receiver's function

The following difference-in-differences model uses Section 4's insights about quasi-random trader entry into the social network to estimate the causal effect of peer performance on $i$'s trading:

$$\begin{aligned} trade.count^{r}_{i,j,t} = \delta_1 join_{i,j,t} + \delta_2 R^{s+}_{-i,j,t} + \delta_3 R^{s-}_{-i,j,t} + \delta_4 join_{i,j,t} \times R^{s+}_{-i,j,t} ... \\ ... + \delta_5 join_{i,j,t} \times R^{s-}_{-i,j,t} + m_t + f_i + b_i + \varepsilon_{i,j,t}. \end{aligned} \tag{5}$$

The dependent variable, $trade.count^{r}_{i,j,t}$, is the number of trades by a trader on the receiving end (superscript $r$) of peer transmission. The independent variable, $join_{i,j,t}$, is equal to one if $i$ has joined myForexBook by time $t$, while $R^{s+}_{-i,j,t}$ $(R^{s-}_{-i,j,t})$[12] is equal to one if traders in peer group $j$ experience positive (negative) contemporaneous returns (senders are given the superscript $s$).[13] In the first set of regressions, $j$ is defined as anyone that has formed a bilateral friendship with $i$ by the end of the sample period and is constant over $t$. Alternative definitions of $j$ are explored in later tests. The model includes trader and week fixed effects,

[12]The subscript $-i$ implies that $i$'s outcome is excluded from the peer-group average.

[13]Indicator variables for positive and negative returns offer a straightforward interpretation of the regression results and also allow us to avoid considerations over the functional form between $i$'s trading and the returns of peer group $j$. However, estimation results are robust to various transformations of $R^{s}_{-i,j,t}$, which can be reproduced upon request.

$f_i$ and $m_t$ respectively. Brokerage fixed effects are $b_i$. Since $b_i$ is collinear to $f_i$, we often interact brokerage dummy variables with $join_{i,j,t}$ to capture any unobservable heterogeneity across brokerages that relates to the timing of entry into myForexBook.

To interpret Eq. 5, $\delta_1$ provides an estimate of the change in $i$'s average weekly trading after joining myForexBook. The contemporaneous correlation between $i$'s trading and the performance of others is represented by $\delta_2$ ($\delta_3$), which accounts for any unobservable shocks that simultaneously affect $j$'s performance and $i$'s trading volume. The interaction between $join_{i,j,t}$ and $R^{s+}_{-i,j,t}$ ($R^{s-}_{-i,j,t}$) examines how the effect of joining the network varies with respect to the contemporaneous performance of peers. Since traders are unable to access the social network's features prior to joining myForexBook, the coefficient on the interaction term $\delta_4$ ($\delta_5$) isolates the causal influence of social interaction.

The model is estimated using OLS to provide a convenient interpretation of the coefficient estimates. However, the results are robust to estimation methods, such as a Poisson model, that are well-suited to handle count data as a dependent variable (available upon request). Furthermore, standard errors are double-clustered by trader and week.

### *5.1. Estimation results of the Receiver's function*

Table 3 presents estimates of Eq. 5. A variety of specifications demonstrate the robustness of our identifying assumptions. First, however, Column I restricts the sample to observations that occur after the trader has joined the social network ($join_{i,j,t} = 1$), thereby offering a reduced-form view of the correlation between trading and the observable returns of peers. The coefficient estimate is equal to 4.15 (s.e. $= 0.9$) for $join_{i,j,t} \times R^{s+}_{-i,j,t}$, which suggests that individuals trade more when the good performance of peers is observable. On the other hand, negative peer returns are uncorrelated with $i$'s trading. The coefficient estimate on $join_{i,j,t} \times R^{s-}_{-i,j,t}$ is 0.50 (s.e. $= 0.5$). To understand how factors unrelated to social interaction shape the correlation between $i$'s trading volume and the returns of $i$'s peers, Column II restricts the sample to observations occurring prior to $i$'s entry into myForexBook ($join_{i,j,t} = 0$). The coefficient estimates on $R^{s+}_{-i,j,t}$ and $R^{s-}_{-i,j,t}$ are not statistically different from zero.

Column III employs a difference-in-differences model, and the coefficients $\delta_4$ and $\delta_5$ provide estimates of the causal effect of social interaction on trading activity. Absent $b_i$, $\delta_4$ is estimated to be 5.40 (s.e. $= 2.0$) and $\delta_5$ is 0.71 (s.e. $= 2.1$). Since $\delta_5$ is not statistically different from zero, the model implies that the poor performance of others does not deter trading. This result could seem surprising because the poor performance of other traders is a credible signal about the degree to which retail participants fail to find success trading in risky markets. Meanwhile, when peers perform well, the average trader issues about five additional trades per week, which amounts to a 20% increase in trading volume. To place the magnitude in perspective, a standard deviation increase in the daily transaction count on the NYSE in 2013 is roughly equal to a 15% increase above the mean.

The coefficient estimates are similar when brokerage-specific factors are considered. Column IV includes brokerage fixed effects instead of trader fixed effects, while Column V interacts $join_{i,j,t}$ with brokerage dummies to allow the average effect of joining myForexBook to vary across groups of traders who belong to the same brokerage. Aggregate market characteristics also have little effect on the coefficients of interest. Columns VI interacts $DXY_t$ with trader dummy variables, while Column VII uses an interaction with $CVIX_t$.

In the last of our initial tests, Column IX includes a triple interaction of $received.message_{i,j,t}$ – an indicator variable equal to one if $i$ receives at least one user message in week $t$ – with $join_{i,j,t}$ and $R^{s+}_{-i,j,t}$ ($R^{s-}_{-i,j,t}$). Column X interacts $average.chatter_t$ with $join_{i,j,t}$ and $R^{s+}_{-i,j,t}$

($R^{s-}_{-i,j,t}$). The coefficient estimate on the triple interaction for positive peer performance is 3.14 (s.e = 0.9) in the former and 3.62 (s.e. = 0.6) in the latter. Thus, the marginal effect of social interaction increases when conversation heightens, which provides evidence that peer-to-peer communication helps explain our findings.

As new traders enter the network, the composition of the control group – traders yet to join the network – changes. Table 4 provides a robust set of alternative specifications that account for ways in which this feature of the data may obscure a causal interpretation of the regression coefficients. Row (1) includes an indicator variable for groups of brokerages that agree with myForexBook during the same month. This specification isolates traders into cohorts of new entrants and is akin to a set of localized regressions, thereby alleviating concerns that brokerages that agree with myForexBook at different points in the sample contain traders with dissimilar unobservable characteristics. For similar reasons, Row (2) restricts the analysis to just the months before and after $i$ joins the network. Row (3) excludes observations that occur before $i$ joins myForexBook and after his or her brokerage agrees with myForexBook. This provides a more rigorous treatment of the control group. Row (4) restricts the observations to traders that join myForexBook in the month that his or her brokerage agrees with myForexBook. This specification accounts for the possibility that traders who join shortly thereafter were the most constrained from engaging in social interaction. All four alternative specifications support our causal conclusions.

Table 4 also explores different specifications for $R^{s}_{-i,j,t}$ and different compositions of $j$, thereby examining the mechanism by which peer influence takes hold. Row (5) uses the one week lag of returns, $R^{s+}_{-i,j,t-1}$ ($R^{s-}_{-i,j,t-1}$), to assess how the past returns of other traders affects future trading. The estimation results are supportive, albeit the estimated effect of positive peer returns falls slightly to around three additional trades.

Specifications (6) and (7) provide additional evidence that peer-to-peer communication is important for peer effects. Row (6) calculates $R^{s+}_{-i,j,t}$ and $R^{s-}_{-i,j,t}$ as the average returns of every peer $j$ who communicates with $i$ via myForexBook's messaging system in week $t$.[14] Similarly, Row (7) includes the returns of all $j$ who have sent at least one user message to $i$ during the sample. The coefficient on $R^{s+}_{-i,j,t}$ is positive and statistically significant at the five percent level, and it is about a third greater in magnitude than previous estimates that use all friends in $j$.

Some peers exert a stronger influence than others. Row (8) restricts the calculation of $R^{s+}_{-i,j,t}$ ($R^{s-}_{-i,j,t}$) to the peer who send the most messages to $i$ over the course of the sample.[15] Also, first impressions have the strongest impact. Row (9) restricts $R^{s+}_{-i,j,t}$ ($R^{s-}_{-i,j,t}$) to the returns of traders in $j$ with whom $i$ forms a bilateral friendship during the first month after joining myForexBook. In both rows (8) and (9), the coefficient on the interaction term $join_{i,j,t} \times R^{s+}_{-i,j,t}$ is around seven and statistically significant at the five percent error level.

## *5.2. A placebo test*

A placebo exercise, which is outlined in Fig. 5 (Panel I), helps determine the extent to which secular trends in the data produce a spurious peer-effect relationship. The estimation is the same as in Eq. 5, but $join_{i,j,t}$ is replaced with a false date of entry that occurs prior to $i$ actually joining myForexBook. The placebo variable, $placebo.join_{i,j,t}$, is equal to one in all weeks following the false date of entry and is zero otherwise. In conducting the test, the first false date of entry is the week before $i$ actually joins. We then roll back the start date for $placebo.join_{i,j,t}$ one week at a time and conduct 20 total regressions. All observations that

[14] When $join_{i,j,t} = 0$, $R^{s}_{-i,j,t}$ is the returns of all traders that message $i$ at any time $t$

[15] When more than one trader $j$ ties for having sent the most messages to $i$, we use the average of those traders' returns.

occur after $i$ actually joins myForexBook are excluded in order to remove any residual effect of the social network from the analysis.

Figure 5 (Panel II) presents kernel density estimates of the distribution of t-statistics from regressions that use false dates for $i$'s entrance into the network. The placebo exercise produces false-positive results with t-statistics above 1.96 in just one out of twenty regressions. Since $i$ is unable to use the social networking features of myForexBook prior to joining, these results suggest that the empirical specification in the previous section has isolated the influence of social interaction on investor trading from other unobservable factors.

All told, the empirical evidence leaves little room for alternative interpretations: individual investors trade more in response to the positive returns of others, while negative peer returns are ignored.

# 6 The Sender's function

Traders respond to good peer performance, a finding that has important implications if positive outcomes are more likely to be transmitted. With an interest in studying revealed attempts to foster communication, Table 5 calculates the probability of sending at least one personal message to any trader in a given week within each quartile of $R_{i,t}$.

There is a positive relation between portfolio performance and contacting others that increases by about six to seven percentage points between the bottom and top quartile of returns (Column (1)). The trend is similar after conditioning the data on a number of observable characteristics. Columns (2) and (3) partition the sample by low and high experience, 0 - 3 and 4+ years, respectively. Columns (4) and (5) divide the data into U.S. and European traders. Columns (6) and (7) divide the data into above and below the median trader age, 36. The final pair of columns, (8) and (9), separates traders into those who practice technical trading strategies versus all others approaches to trading.

## *6.1. An empirical model of the Sender's function*

We use the method of partially overlapping peer groups to formally test if good performance causes a trader to contact others (Bramoulle, Djebbari, and Fortin (2009) and De Giorgi, Pellizzari, and Redaelli (2010)). A difference-in-differences framework cannot be used to estimate the Sender's function, because myForexBook traders cannot communicate with one another until after joining the network. Hence, the outcome variable – $sent.message_{i,j,t}$, an indicator equal to one if $i$ sends at least one message to another trader in week $t$ – is always equal to zero prior to network entry. However, the overlapping-peers methodology also utilizes the dynamics in trader entry to alleviate concerns that observing the actions of $i$'s peers simultaneously influences $i$'s behavior, or what is referred to as "reflection" (Manski (1993)).

To describe the estimation technique, we borrow Manski's (1993) terminology, in which the standard peer-effects model has a linear-in-means setting in which the independent variable is $y_{i,j,t}$ for trader $i$ in peer group $j$. The right-hand side variables include average peer group outcomes, $\bar{y}_{-i,j,t}$, average characteristics of the peer-group, $\bar{X}_{-i,j,t}$, average own characteristics, $X_{i,j,t}$, and a peer-group fixed-effect variable, $g_{j,t}$. The model contains three types of peer effects: *endogenous effects*, $\bar{y}_{-i,j,t}$, *exogenous effects*, $\bar{X}_{-i,j,t}$, and *correlated effects*, $g_{j,t}$. The peer-effects model is therefore:

$$y_{i,j,t} = \alpha + \beta \cdot \bar{y}_{-i,j,t} + \gamma \cdot \bar{X}_{-i,j,t} + \delta \cdot X_{i,j,t} + \zeta \cdot g_{j,t} + \epsilon_{i,j,t}. \tag{6}$$

Bramoulle, Djebbari, and Fortin (2009) and De Giorgi, Pellizzari, and Redaelli (2010) prove that Eq. 6 allows endogenous and exogenous peer effects to be separately identified under the reasonable assumption that the links are intrasensitive – that is, traders can be connected to a common peer without necessarily being connected to each other. An example from De Giorgi, Pellizzari, and Redaelli (2010) provides intuition.[16] If traders $A$ and $B$ are connected to each other, but only $A$ is connected to $C$, peer effects are $y_A = f(B, C)$, but for $B$ and $C$ they are equal to $f(A)$. This reduces the model to a just-identified set of linear equations, one for each of $N$ individuals, which isolates each individual's contribution to peer effects.

The identifying assumption is reasonable within our setting, because as outlined in Section 4, agreements between myForexBook and participating brokerages predict new membership. The process of trader entry causes peer characteristics – $\bar{y}_{-i,j,t}$, $\bar{X}_{-i,j,t}$, and $g_{j,t}$ – to evolve on a weekly basis in a manner that is plausibly exogenous. This causes trader $i$'s peer group $j$ to have a unique composition (partially overlapping). The model can thus extract the contribution of each new peer without concerns over time-invariant confounds.

Within the context of the myForexBook setting, the partially overlapping peers model is

$$\begin{aligned} sent.message_{i,j,t} = \beta_1 R_{i,j,t} + \beta_2 peer.\overline{R}_{-i,j,t} + \beta_3 log.peer.\overline{messages}_{-i,j,t} + \beta_4 \bar{X}_{-i,j,t} ... \\ ... + \beta_5 X_{i,j,t} + m_t + f_i + g_{j,t} + \epsilon_{i,j,t} \qquad (7) \end{aligned}$$

and can be estimated using OLS.[17] The independent variable of interest, $i$'s portfolio returns, $R_{i,j,t}$, is normalized to standard deviations about the mean to aid in interpretation. Endogenous peer effects are captured by $peer.\overline{R}_{-i,j,t}$, the size-weighted average of returns

[16]Lewellen (2013) also provides an example of the estimation technique used in the context of executive compensation.

[17]The results are robust to the use of a logistic model, but OLS offers a convenient interpretation of the coefficient estimates.

for traders in $j$, which we also normalize. Hence, the linear model can be thought of as capturing the effect of $i$'s returns benchmarked by the returns of $i$'s peers. Endogenous peer effects also include $log.peer.\overline{messages}_{-i,j,t}$, which is equal to the natural logarithm of one plus the average number of messages sent by members of $j$ at time $t$. Exogenous peer effects, $\bar{X}_{-i,j,t}$, include the characteristics of traders in $j$, for example the average age of $i$'s friends. Exogenous own characteristics are represented by $X_{i,j,t}$, but tend to be absorbed by trader fixed effects, $f_i$. Including binary variables for membership in each $i$'s peer-group would be computationally intractable. Thus, to approximate $g_{j,t}$, we use the topology of the network and a computational network algorithm – outlined in Girvan and Newman (2002) – to sort nearby traders into "communities", which are included in the regression analysis as a series of dummy variables.

The coefficient of interest, $\beta_1$, is the estimated change in the probability of sending a message for a one standard deviation increase in $i$'s returns. The coefficients on the right-hand-side peer-effects variables can be interpreted similarly. For example, the addition of new traders to $j$ over time allows the model to extract the influence of a one-unit increase in the average age of $i$'s peer group on the probability that $i$ sends a user message.

### *6.2. Estimation results of the Sender's function*

Table 6 presents coefficient estimates from the peer-effects model in Eq. 7. Recall, the estimation is restricted to weeks $t$ after $i$ joins myForexBook, because traders can only send messages after joining, which reduces the number of observations to 66,418. In Columns I through V, peer group $j$ is updated weekly to include traders who have joined myForexBook and formed a bilateral friendship with $i$ by week $t$. Therefore, $j$ is called *restricted.peers* because there are other traders in myForexBook who are not connected to $i$.

Estimates of Eq. 6 provide evidence that good returns cause message sending. The coefficient estimate for $\beta_1$ is 0.016 (s.e = 0.007) (Column I). Column II controls for the size of $i$'s peer group and Column III includes $g_{i,j}$. Column IV adds endogenous peer effects to the right-hand side, while Column V adds exogenous peer characteristics. The coefficient estimate for $\beta_1$ remains positive (roughly 0.015) and statistically significant at the five percent error level. In unreported analysis, our estimation results are robust to the inclusion of $DXY_t$ and $CVIX_t$ interacted with trader dummy variables.

The estimation results are economically meaningful. An increase from the 10th to the 90th percentile of the distribution of weekly returns increases the probability of contacting other traders by about five percentage points. The magnitude of the increase is roughly equivalent to a 25% increase above the average unconditional probability of issuing communications. Moreover, the other right-hand-side variables in the peer-effects model do not appear to predict communications, which suggests that among plausible observable variables, traders are motivated to communicate mostly by their own returns.

Alternative treatments of $i$'s peer group provide evidence that the model has successfully handled the reflection problem. Columns VI and VII replace $i$'s peer-group $j$ with $eventual.peers$, traders who have yet to join myForexBook as of $t$, but become friends with $i$ after joining. Similarly, Columns VIII and IX replace $j$ with $excluded.peers$, all traders who have joined myForexBook by $t$, but who never befriend $i$. Recategorizing $j$ is similar to a placebo exercise, helping determine the extent to which the observable behavior of $i$'s connections confound the estimates of $\beta_1$. The group of $eventual.peers$ is especially useful for this purpose, because they have characteristics that are later revealed to be compatible with $i$, but the dynamics of entrance into myForexBook prevents $i$ from being connected to $j$ at time $t$. Therefore, $eventual.peers$ and $restricted.peers$ are likely to have a similar unobservable relationship with $i$. However, the groups differ in the ability to be observed by $i$. Indeed, our estimates of $\beta_1$ are robust to these alternatives with little change in the coefficient value or statistical significance.

Appendix A.4 provides a discussion of some additional robustness checks, which include Granger-causality tests as well as attempts to instrument for portfolio returns. To summarize, returns Granger-cause message sending, while message sending does not Granger-cause returns. Candidate instruments produce a positive relation between returns and message sending, but it is difficult to reliably predict returns in the first-stage, meaning that we tend to have weak instruments. In closing, it appears as if traders prefer to communicate following investment successes.

# 7 Population dynamics

The Han and Hirshleifer (2013) model predicts that a positive shock to trader social interaction increases conversation to Active strategies so long as the Sender and Receiver functions are positively sloped. The development of online social networks that allow traders to better communicate about their investments can be viewed as a shock to baseline sociability.

Aggregate trends in the myForexBook data provide suggestive evidence that the social network acts as a propagation mechanism that supports the growth of active investing. In particular, since the inception of myForexBook, average trading and the volatility of portfolio returns both increase gradually. Among traders who have joined myForexBook by week $t$, the average number of trades per week increases from around 20 at the start of the sample to around 30 near the end (Fig. 6). Meanwhile, the average weekly trading of those outside the network remains close to 20 at both ends of the sample. Despite the divergence in averages, there is considerable comovement between the two groups with a Pearson's correlation coefficient of around 70%, which is consistent with the notion of treatment and control group.

Also indicative of the propagation of active strategies, the average volatility of portfolio returns among myForexBook traders increases by as much as 25% over the sample period (Fig. 7). Meanwhile, the portfolio volatility of traders outside the social network is similar at both ends of the sample period, even though the two groups experience similar fluctuations over time.

## *7.1. Supporting evidence and alternative explanations*

The trends toward active investing tend to be related to social interaction, rather than contemporaneous factors such as changes in market characteristics. First, the channels of communication within the network appear to favor Active traders establishing a central location and encouraging Passive traders to increase their trading intensity. For illustrative purposes, we divide traders into two groups and call traders with the most (least) trading "Active" ("Passive").[18] Fifty-three percent of all friendships involve the pairing of an Active with a Passive trader. Active/Passive pairings would form 45.7% of all friendships if pairs of traders were drawn at random. This is consistent with the notion that strategies transmit from Active to Passive trader, rather than communication taking place mostly among similar pairs of traders.

Furthermore, consistent with an increase in strategy transmission, the frequency with which traders use the social network has increased over the sample period. In August 2009, myForexBook began to record the number of logins to the website, potentially an important

[18] We restrict the Active group according to two criteria: (1) total trades by an individual must exceed the median, (2) and the frequency with which they trade during a given week must also exceed the median.

proxy for social network usage. The median number of logins per trader has roughly doubled over the course of the sample from around five to close to ten times per month.

Changes in the forex market also appear unlikely to explain our findings. Traders could experience more volatile returns if currency markets become more volatile. Contrary to this argument, implied volatility, $CVIX_t$, declines by about a third between the beginning of 2009 and the end of 2010. Likewise, the average value of the U.S. dollar relative to other major currencies, $DXY_t$, is similar at both ends of our sample.

Lastly, despite a noticeable increase in trading frequency, average trader profitability is not statistically different between the start and end of the sample. Therefore, consistent with the prediction that social interaction contributes to the active-investing puzzle, myForexBook users trade more and have more volatile portfolios, but are not compensated with higher profits.

# 8 Concluding remarks

We analyze new data on retail traders who participate in an online social network and provide support for the hypothesis that social interaction promotes active trading. In doing so, we are among the first to examine peer-to-peer communication between traders and to document a few novel empirical findings. Individuals trade more when their peers perform well, but do not respond to losses incurred by peers. They also tend to contact other traders following their own investment successes, even though it could be equally beneficial to notify the community about potential losses. Overall, several pieces of empirical evidence suggest the social network has contributed to the propagation of Active strategies, which is consistent with Han and Hirshleifer's (2013) population evolution model, but is difficult to reconcile with canonical models of heterogeneous information aggregation (Grossman (1976) and Hellwig (1980), among others).

Despite similarities between the traders in the myForexBook data and those in other studies of retail stock market participants, it is possible that our results are most applicable to socially susceptible traders. Since it is beyond the scope of our data to address this critique, we direct readers towards contemporaneous research. Using data from a representative household survey and an approach similar to Hong, Kubik, and Stein (2004), Heimer (2014) shows that social individuals are more likely to exhibit higher portfolio turnover conditional on asset market participation. The relation is most pronounced among individuals with observable characteristics that – according to established research – make them more likely to trade actively. Heimer (2014) lacks comparable detail in investor portfolios and social activity, or the ability to make causal claims, but the results suggest that this paper's findings apply more broadly.

Lastly, we consider our findings to be timely considering the unprecedented growth in technologies over the past decade that have lowered the cost and hastened the pace of communication. As of this writing, we know of close to ten new English-language trading social networks that have developed since we acquired myForexBook's data. The most prominent discount brokerages in the U.S. have also revealed an awareness of this growing industry that goes back at least as far as E*Trade Financial Corporation's year 2000 purchase of the instant messaging service Cahoots, Inc. Finance theorists have taken notice with a renewed interest in modeling information transmission in networks (Han and Yang (2013), Walden (2013), and Andrei and Cujean (2014)). This is certainly a worthy endeavor, because social interaction is an important driver of investor behavior.

# Appendix

## *A.1. Supplementary algebra*

In this section, we present the algebra used to generate the result in Eq. 13. With the intention of guiding this paper's reader, we include a direct recapitulation of some important pieces of Han and Hirshleifer (2013). Curious readers should consult the Han and Hirshleifer (2013) manuscript for a more thorough description.

Suppose there are $n$ total traders. The fraction of traders of a given type is $f_i \equiv \frac{n_i}{n}$, where $f_A + f_P = 1$ and $f'_A + f'_P = 1$, with the apostrophe indicating the following period. The change in $f_A$ is

$$f'_A - f_A = \begin{cases} \frac{1}{n} & w/probability \;\; (\chi/2)\, T_{AP}\,(R_A) \\ -\frac{1}{n} & w/probability \;\; (\chi/2)\, T_{PA}\,(R_P) \\ 0 & w/probability \; 1 - (\chi/2)\,[T_{AP}\,(R_A) + T_{PA}\,(R_P)] \end{cases} \tag{8}$$

where $\chi$ is the probability that a pair of $A$ and $P$ traders is drawn at random,

$$\begin{aligned} \chi &\equiv \left(\frac{n_A}{n}\right)\left(\frac{n-n_A}{n-1}\right) + \left(\frac{n-n_A}{n}\right)\left(\frac{n_A}{n-1}\right) \\ &= \frac{2nf\,(1-f)}{(n-1)}. \end{aligned} \tag{9}$$

Returns to $A$ and $P$ strategies consist of a common component, $r$, (for example, the market portfolio) and an idiosyncratic component, $\varepsilon_i$, both random variables with mean zero that are mutually independent. However, $A$ and $P$ strategies differ in their sensitivity to a common factor, $\beta_A > \beta_p \geq 0$, and the variance of the idiosyncratic component, $\sigma_A^2 > \sigma_P^2$.[19] Realized returns, $R_i$, are a linear combination of $\beta_i$ and $\varepsilon_i$, and there is a penalty, $D$, to using Active strategies which may reflect trading costs, opportunity costs, or a lack of diversification. Thus, $R_A = \beta_A r + \varepsilon_A - D$ and $R_P = \beta_P r + \varepsilon_P$.

According to Han and Hirshleifer (2013), the strategy transmission probability, $T$, from Sender $A$ to Receiver $P$ is

$$T_{AP}(R_A) = r(R_A)s(R_A) \tag{10}$$

and from Sender $P$ to Receiver $A$,

$$T_{PA}(R_P) = r(R_P)s(R_P) \tag{11}$$

Han and Hirshleifer (2013) substitute the Sender and Receiver functions, Equations 1 and 2, into Equations 10 and 11 and subtract,

$$\begin{aligned} T_{AP}(R_A) - T_{PA}(R_P) =& a\lambda\left(R_A^3 - R_P^3\right) + B\left(R_A^2 - R_P^2\right) + C\left(R_A - R_P\right) \\ =& a\lambda\left[\left(\beta_A^3 - \beta_P^3\right)r^3 + 3r^2\left(\beta_A^2\varepsilon_A - \beta_P^2\varepsilon_P\right) + 3r\left(\beta_A\varepsilon_A^2 - \beta_P\varepsilon_P^2\right) + \varepsilon_A^3 - \varepsilon_P^3\right] \\ &+ B\left[\left(\beta_A^2 - \beta_P^2\right)r^2 + 2r\left(\beta_A\varepsilon_A - \beta_P\varepsilon_P\right) + \varepsilon_A^2 - \varepsilon_P^2\right] + C\left[\left(\beta_A - \beta_P\right)r + \varepsilon_A - \varepsilon_P\right] \\ &+ D\left\{-\left(r\beta_A + \varepsilon_A\right)\left[3a\beta\left(r\beta_A + \varepsilon_A\right) + 2B\right] - C\right\} + D^2\left[3a\beta\left(r\beta_A + \varepsilon_A\right) + B\right] \\ &- aD^3\lambda. \end{aligned} \tag{12}$$

where $B = a\gamma + b\lambda$ and $C = b\gamma + c\lambda$.

Given a fixed population of traders and assuming that the returns to the strategies have zero skewness, the expected unconditional change in the fraction, $f_A$, of $A$ traders from Eq.

[19] A few unreported tests support the variance assumptions in the myForexBook data.

12 is:

$$\begin{aligned}\left(\frac{2n}{\chi}\right)E\left[\triangle f_A\right] =& E\left[T_{AP}(R_A)\right] - E\left[T_{PA}(R_P)\right] \\ =& B\left((\beta_A^2 - \beta_P^2)\sigma_r^2 + (\sigma_A^2 - \sigma_P^2)\right) + a\lambda\left(-3\sigma_A^2 - D^2 - 3\sigma_r^2\beta_A^2\right)D + BD^2 - CD \end{aligned} \tag{13}$$

Taking the partial derivative with respect to $\gamma$ produces the result in Eq. 3.

## *A.2. A comparison of myForexBook traders to Finnish traders*

We address the possibility that myForexBook traders are different than the typical retail trader. Traders who join myForexbook are potentially more susceptible to social influences than the population of retail investors, contaminating our estimates of the Sender and Receiver functions. In order to test the susceptibility of traders to social influences, we replicate the epidemic model of Shive (2010), where the probability of a trader opening and closing a position in a certain security depends on the product of the number of the people who made the same trade previously (infected) and the number of people who did not (susceptible). The factor measures the likelihood that two people of different opinions about the asset will meet and one person will influence the other. The coefficient on the factor therefore measures the sensitivity of individual investor trades to social influence. The Shive (2010) study uses the population of traders in Finland. Therefore, it is reasonable to suggest that our results are unbiased if myForexBook traders behave similarly.

We divide the period March to September 2010 into 60 minute intervals and run the following four logistic regressions:

$$logit(SocialTrade_{p,t}) = \beta_0 + \beta_1 SI_{p,t} + \beta_2 Controls_{p,t} + \varepsilon_{i,t}$$

The dependent variable,

$$SocialTrade_{p,t} = \{OpenLong_{p,t}, OpenShort_{p,t}, CloseLong_{p,t}, CloseShort_{p,t}\}$$

is equal to one if the trader opened (or closed) a long (or short) position in the currency pair $p$ in interval $t$ and zero otherwise. $SI$ is the proportion of friends who had a position in the currency pair at the beginning of the 60 minute interval. Controls include an indicator variable to show if there are macroeconomic news announcements in the previous 60 minutes, the previous 60 minutes' return on the currency pair, and the proportion of traders within the entire population who had a position in the currency pair. The estimation is performed separately for each trading day. We estimate the equations for the three most popular currency pairs (EUR/USD, GBP/USD, USD/JPY), together comprising around 57% of total trading volume.

The estimation results, presented in Table A.1, imply that myForexBook traders are similarly susceptible to social influences as are the population Finnish investors. Shive (2010) finds a positive coefficient on community ownership 98% of the time, while our estimation results varying by currency show positive coefficients only about two-thirds of the time.

Table A.1: **Epidemic Effects in Trading**

The following table presents estimates of the following logistic regression:

$$SocialTrade_{c,t} = \beta_0 + \beta_1 SI_{c,t} + \beta_2 Controls_{c,t} + \varepsilon_{i,t}$$

where,

$$SocialTrade_{c,t} = \{OpenLong_{c,t}, OpenShort_{c,t}, CloseLong_{c,t}, CloseShort_{c,t}\}$$

The dependent variable is 1 if the trader opened (or closed) a long (or short) position in the currency pair $c$ in interval $t$ and zero otherwise, while $SI$ is the proportion of friends who had a position in the currency pair at the beginning of the 60 minute interval $t$. Controls include an indicator variable to show if there are macroeconomic news announcements in the previous 60 minutes, the previous 60 minutes' return on the currency pair, and the proportion of traders within the entire population who had a position in the currency pair. Average odds-ratios by currency pair and average standard errors are displayed. Percentage of positive coefficients are next to estimates. The equation is estimated daily across 123 trading days. Only traders who are members of myForexBook at the 12PM GMT at the beginning of the day are considered in the estimation.

| | $OpenLong$ | | $OpenShort$ | | $CloseLong$ | | $CloseShort$ | |
|---|---|---|---|---|---|---|---|---|
| | $SI_{i,t}$ | $\% \, SI_{i,t} > 0$ | $SI_{i,t}$ | $\% \, SI_{i,t} > 0$ | $SI_{i,t}$ | $\% \, SI_{i,t} > 0$ | $SI_{i,t}$ | $\% \, SI_{i,t} > 0$ |
| $EUR/USD$ | 1.002** | 81.18% | 1.001** | 69.11% | 1.002* | 56.91% | 1.001* | 56.91% |
| | (0.49) | | (0.49) | | (0.59) | | (0.56) | |
| $GBP/USD$ | 1.003** | 57.72% | 1.002* | 61.78% | 1.000 | 52.03% | 1.001 | 53.65% |
| | (0.49) | | (0.50) | | (0.70) | | (0.65) | |
| $USD/JPY$ | 1.002** | 81.30% | 1.001** | 73.98% | 1.001* | 56.09% | 1.001* | 55.18% |
| | (0.47) | | (0.43) | | (0.63) | | (0.56) | |
| $N = 136,632$ for each daily regression | | | | | | | | |

Odds-ratios; Mean of standard errors in parentheses

* $p < 0.10$ , ** $p < 0.05$ , *** $p < 0.01$

### *A.3. Identifying assumptions: Empirical and anecdotal evidence*

Table 2 provides evidence in support of our identifying assumption, that the agreements between a brokerage and myForexBook are uncorrelated with trader characteristics. The data is partitioned into the first set of traders that join myForexBook and the last to join, which are called $first_i$ and $last_i$, respectively. Panel I provides a set of difference in means tests to compare trader characteristics in the two groups. The first set of columns compares the first and last 250 traders to join, the second set compares the first and last 500, while the third set compares the first and last 1,000. There appears to be little difference in the observable trader characteristics between the first and last to join, because few variables have t-statistics that are above a threshold for statistical signficance.

More formally, Panel II presents a balancing test, which is equivalent to estimating a Probit model with the dependent variable being an indicator equal to one for traders who are among the first to join myForexBook. Using the observable characteristics in Panel I as independent variables, the estimation produces a poor model fit with a pseudo-$R^2$ of around 0.01 or less in all three regressions – using the first and last 250, 500, and 1,000 to join, respectively. Furthermore, out of 36 potential covariates, only one coefficient is statistically significant at at least the ten percent error level.

These findings broadly support the notion that trader selection onto participating brokerages is not a concern. However, we offer some additional anecdotal evidence. First, while it is difficult for us to quantify, traders have a limited menu of brokerages to choose from. Different brokerages have different capital requirement that usually ranges from a few thousand to tens-of-thousands of U.S. dollars. There are also substantial regulatory differences across countries, and since all of the brokerages require new clients to provide proof of citizenship, it would for example, be impossible for a U.S. trader to leave their U.S. brokerage for one based in Germany with the intention of joining myForexBook. Secondly, there are considerable opportunity and potential monetary costs associated with changing brokerages. When a trader closes all of their positions and transfers money to a new brokerage, the trader potentially loses an opportunity to trade during that time, and likely suffers from order execution risks when withdrawing funds. Moreover, different brokerages have different software that traders must learn and different features such as the length and precision of historical price-feeds. Therefore, it is difficult to envision how these factors would potentially relate to the timing of agreements between participating brokerages and myForexBook, which lends additional confidence in our identification strategy.

### *A.4. Robustness of the Sender's function and other considerations*

The positive relation between returns and message sending is robust to several additional considerations. Reverse causality seems unlikely to obscure a causal interpretation because the empirical technology isolates the idiosyncratic effect of trader returns from any peer-group confounds and the myForexBook network is unlikely to have sufficient market power to influence forex prices. Regardless, a Granger-causality test between sending a message and returns in the next period provides empirical evidence. Returns Granger-cause sending a message, while sending a message does not Granger-cause returns (Table A.2).

A few potential instruments for Sender returns offer corroborative results (available upon request). Using the daily change in $DXY_t$ and an implied currency volatility index, $CVIX_t$, as instrumental variables in a two-staged least-squares analysis produces similar second-stage coefficient estimates, but prove to be weak instruments (F-values between three and five). We also use the surprise component of macroeconomic news releases provided by Bloomberg to forecast individual investor returns. However, the first-stage estimates are unreliable because macro variables tend to be poor predictors of exchange rate movements at horizons shorter than a year. Lastly, we use trader account balances as a proxy for wealth and a subsequent predictor of returns, because there is evidence that wealthier investors utilize more productive search efforts (Bonaparte and Fabozzi (2011)). Indeed, weekly account balances are positively correlated with subsequent returns (a first-stage F-value of around 90). While the second-stage estimates produce a relation between returns and messaging that is similar to prior regressions, a Durbin-Hausman-Wu test indicates that OLS produces more efficient estimates. Moreover, the quasi-random propagation of peer-groups utilized in Eq. 7 likely provides a more satisfying treatment of the necessary exclusionary restrictions than the use of account balances as an instrument.

Table A.2: **Granger-Causality Test for Returns and Message Sending**

This table reports the results of the linear Granger causality test on returns and sending messages. $L_{returns}$ and $L_{messages}$ denotes the number of backwards lags on dollar returns and number of messages sent, respectively. Lag lengths are set with the Aikike Information Criterion. *Sign* denotes the marginal significance of the computed $\chi^2$ statistics used to test the restrictions implied by the null hypothesis of no Granger causality.

| $H_0$ : returns cause sending messages | | | | | $H_0$ : sending messages cause returns | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $L_{returns}$ | $L_{messages}$ | $\chi^2$ | $Sign$ | $N$ | $L_{returns}$ | $L_{messages}$ | $\chi^2$ | $Sign$ | $N$ |
| 3 | 8 | 22.36 | 0.051 | 38,645 | 6 | 5 | 3.18 | 0.208 | 31,201 |

# References


Andrei, D. and Cujean, J. (2014). Information percolation, momentum, and reversal. *Working Paper*.

Antweiler, W. and Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *Journal of Finance*, 59(3):1259–1294.

Barber, B. M., Lee, Y.-T., Liu, Y.-J., and Odean, T. (2009). Just how much do individual investors lose by trading? *Review of Financial Studies*, 22(2):609–632.

Barber, B. M. 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. M. and Odean, T. (2001a). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, 116(1):261–292.

Barber, B. M. and Odean, T. (2001b). The internet and the investor. *Journal of Economic Perspectives*, 15(1):41–54.

Ben-David, I., Heimer, R. Z., and Hou, K. (2014). Does retail trading improve stock price efficiency? *Federal Reserve Bank of Cleveland Internal Document*.

Bénabou, R. and Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, 117(3):871–915.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., and Milkman, K. L. (2011). The effect of providing peer information on retirement savings decisions. Working Paper 17345, National Bureau of Economic Research.

Biais, B., Hilton, D., Mazurier, K., and Pouget, S. (2005). Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market. *Review of Economic Studies*, 72(2):287–312.

Bildik, R., Ozsoylev, H. N., Walden, J., and Yavuz, M. D. (2014). Investor networks in the stock market. *Review of Financial Studies*, 27(2):1323–1366.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008.

Bonaparte, Y. and Fabozzi, F. (2001). Household search choice: Theory and evidence. *Applied Economics*, 43(26):3835–3847.

Bramoulle, Y., Djebbari, H., and Fortin, B. (2009). Identification of peer effects through social networks. *Journal of Econometrics*, 150(1):41–55.

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

Chen, H., De, P., Hu, Y., and Huang, B.-H. (2014). Wisdom of crowds: The value of stock opinions transmitted through social media. *Review of Financial Studies*, 27:1367–1403.

Choi, J. J., Laibson, D., and Metrick, A. (2002). How does the internet affect trading? Evidence from investor behavior in 401(k) plans. *Journal of Financial Economics*, 64(3):397–421.

Cohen, L., Frazzini, A., and Malloy, C. (2008). The small world of investing: Board connections and mutual fund returns. *Journal of Political Economy*, 116(5):951–979.

De Giorgi, G., Pellizzari, M., and Redaelli, S. (2010). Identification of social interactions through partially overlapping peer groups. *American Economic Journal: Applied Economics*, 2(2):241–75.

Duflo, E. and Saez, E. (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *The Quarterly Journal of Economics*, 118(3):815–842.

Engelberg, J., Gao, P., and Parsons, C. A. (2012). Friends with money. *Journal of Financial Economics*, 103(1):169 – 188.

Engelberg, J. and Parsons, C. A. (2014). Worrying about the stock market: Evidence from hospital admissions. *AFA Annual Meetings Paper*.

Foucault, T., Sraer, D., and Thesmar, D. J. (2011). Individual investors and volatility. *The Journal of Finance*, 66(4):1369–1406.

Fracassi, C. and Tate, G. (2012). External networking and internal firm governance. *The Journal of Finance*, 67(1):153–194.

French, K. R. (2008). Presidential address: The cost of active investing. *Journal of Finance*, 63(4):1537–1573.

Girvan, M. and Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826.

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

Grinblatt, M. and Keloharju, M. (2009). Sensation seeking, overconfidence, and trading activity. *Journal of Finance*, 64(2):549–578.

Grossman, S. (1976). On the efficiency of competitive stock markets where trades have diverse information. *The Journal of Finance*, 31(2):573–585.

Han, B. and Hirshleifer, D. (2013). Self-enhancing transmission bias and active investing. *AFA Annual Meetings Paper*.

Han, B. and Yang, L. (2013). Social networks, information acquisition, and asset prices. *Management Science*, 59(6):1444–1457.

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

Hellwig, M. F. (1980). On the aggregation of information in competitive markets. *Journal of Economic Theory*, 22(3):477–498.

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

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

Hou, K. and Moskowitz, T. J. (2005). Market frictions, price delay, and the cross-section of expected returns. *Review of Financial Studies*, 18(3):981–1020.

Ivković, Z. and 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. and Knüpfer, S. (2012). Peer performance and stock market entry. *Journal of Financial Economics*, 104(2):321–338.

King, M. R. and Rime, D. (2010). The $4 trillion question: What explains FX growth since the 2007 survey? *BIS Quarterly Review*.

Kumar, A. (2009). Who gambles in the stock market? *Journal of Finance*, 64(4):1889–1933.

Lewellen, S. (2013). Executive compensation and peer effects. *Working Paper*.

Li, G. (2014). Information sharing and stock market participation: Evidence from extended families. *The Review of Economics and Statistics*, (forthcoming).

Linnainmaa, J. T. (2010). Do limit orders alter inferences about investor performance and behavior? *The Journal of Finance*, 65(4):1473–1506.

Linnainmaa, J. T. (2011). Why do (some) households trade so much? *Review of Financial Studies*, 24(5):1630–1666.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3):531–42.

Mizrach, B. and Weerts, S. (2009). Experts online: An analysis of trading activity in a public Internet chat room. *Journal of Economic Behavior & Organization*, 70(1-2):266–281.

Pool, V., Stoffman, N., and Yonker, S. (2014). The people in your neighborhood: Social interactions and mutual fund portfolios. *The Journal of Finance*, forthcoming.

Popadak, J. A. (2014). Dividend payments as a response to peer influence. *AFA Annual Meetings Paper*.

Shive, S. (2010). An epidemic model of investor behavior. *Journal of Financial and Quantitative Analysis*, 45(01):169–198.

Shue, K. (2013). Executive networks and firm policies: Evidence from the random assignment of MBA peers. *Review of Financial Studies*, 26(6):1401–1442.

Walden, J. (2014). Trading, profits, and volatility in a dynamic information network model. *AFA Annual Meetings Paper*.

Figure 1: **myForexBook User Profiles**

This figure is an example of a myForexBook user's profile web-page. Any details that could be used to identify the social network or its participants have been removed from the image.

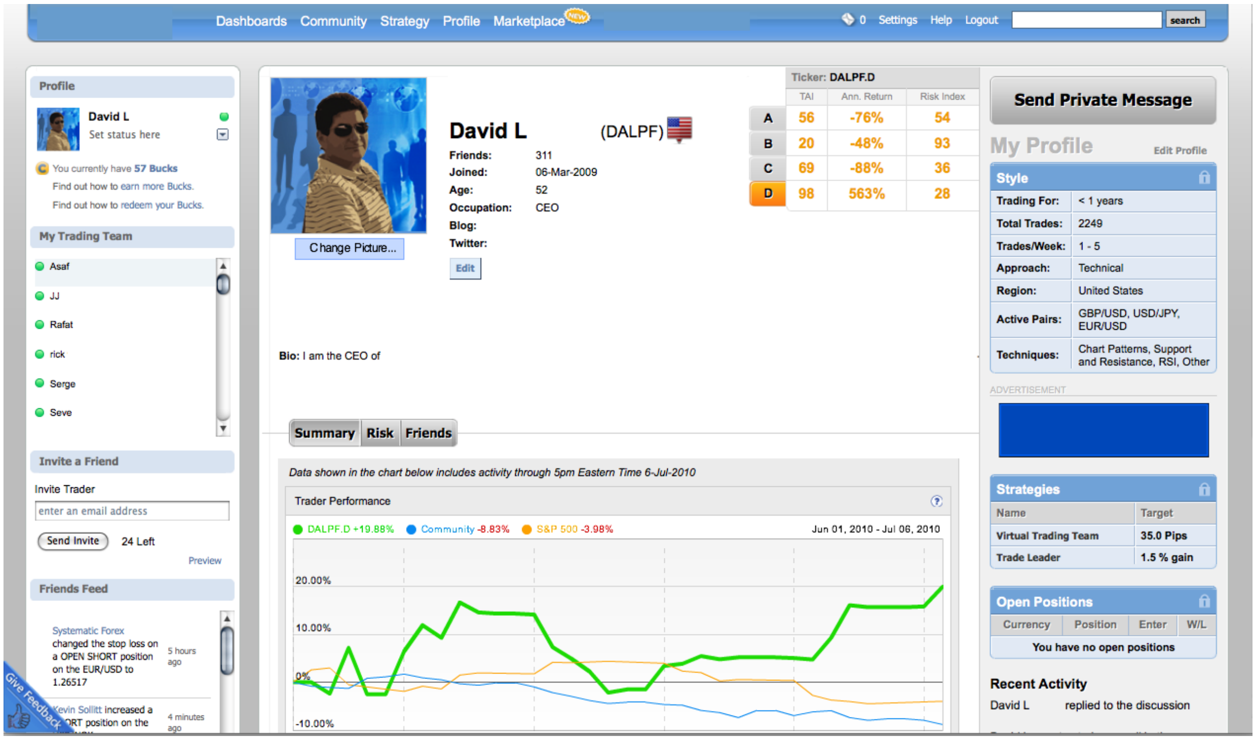


Figure 2: **myForexBook "Dashboard"**

This figure comes from the myForexBook web-platform. After forming a friendship with another trader, the friend's trades are displayed in real-time in the manner demonstrated below.

**Positions**

All **Open** Closed

| currency | user | status | position | time | SL | enter | TP | close time | close price | margin | social indicators |
|---|---|---|---|---|---|---|---|---|---|---|---|
| USD/CHF | Sarah Conors | PENDING | SHORT | 4:09pm | | 1.15400 | | | | | |
| USD/JPY | Sven Olusk | OPEN | SHORT | 2:36pm | 96.5490 | 96.3400 | | | | 40:1 | |
| GBP/USD | Phillip Alford | OPEN | LONG | 12:09pm | 1.45000 | 1.46000 | 1.47500 | | | 7:1 | |
| GBP/USD | Sarah Conors | OPEN | SHORT | 11:18am | 1.47058 | 1.46210 | 1.44788 | | | 7:1 | |
| EUR/USD | Sarah Conors | OPEN | SHORT | 9:21am | 1.32120 | 1.30435 | 1.27120 | | | 7:1 | |
| GBP/USD | Kate Taylor | OPEN | LONG | 8:02am | 1.45000 | 1.46017 | | | | 4:1 | |
| USD/CHF | Thomas Kostek | OPEN | SHORT | 7:38am | 1.15500 | 1.15770 | | | | 6:1 | |
| GBP/USD | Thomas Kostek | OPEN | SHORT | Apr 27 | | 1.45990 | 1.44500 | | | 1:1 | |
| EUR/USD | Phillip Alford | OPEN | LONG | Apr 27 | 1.30500 | 1.30000 | | | | 3:1 | |
| EUR/USD | Thomas Kostek | OPEN | LONG | Apr 27 | 1.29370 | 1.30204 | 1.32870 | | | 33:1 | |

Figure 3: **The Topology of the myForexBook Network**

This figure illustrates the complete set of connections between myForexBook traders at the end of the sample period. We use Blondel, et al (2008))'s visualization algorithm and the network software Gephi to show that certain groups of traders are more tightly connected to each other than they are to others, which is represented by an arbitrary color spectrum from dark red to purple. The size of each node is roughly proportional to the number of friends each trader has.

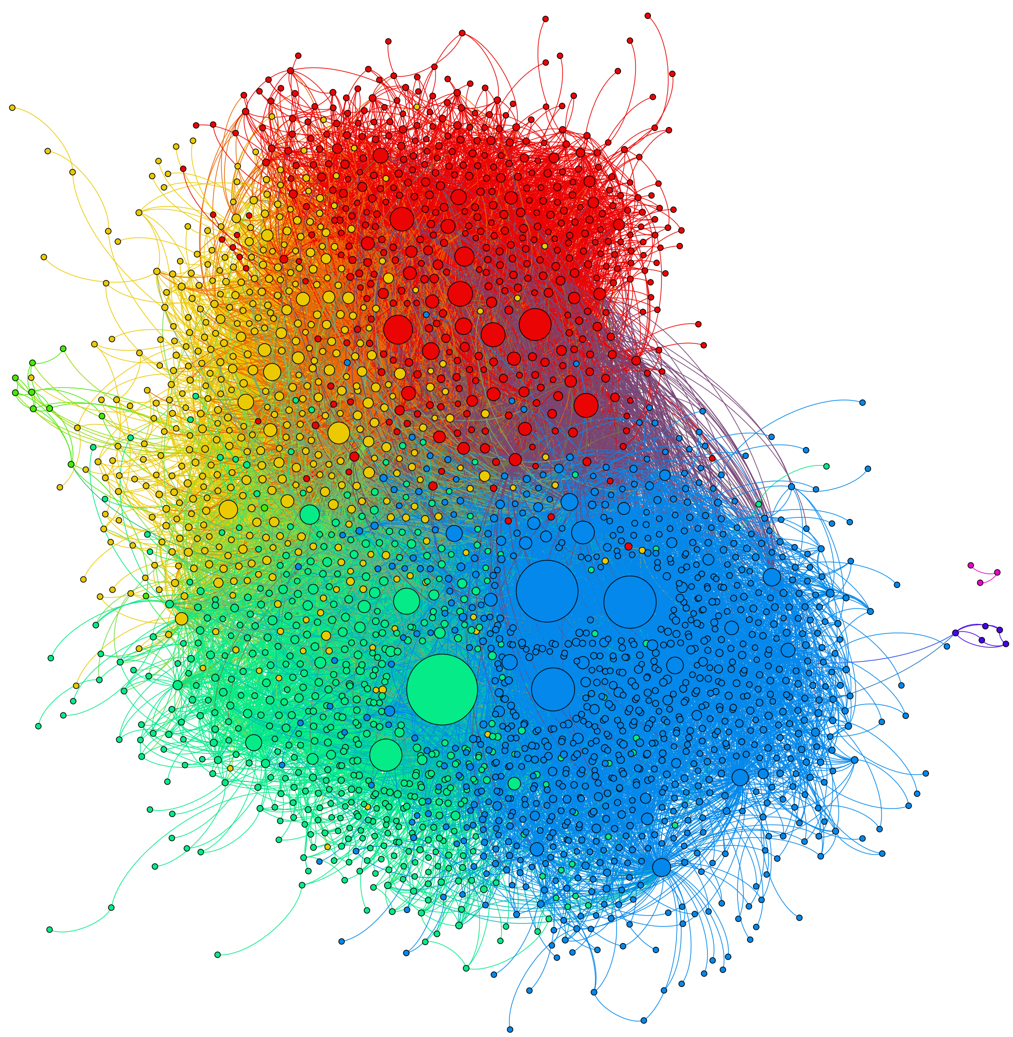


Figure 4: **New Partnerships Between myForexBook and Retail Brokerages**

This figure illustrates the formation of partnerships between myForexBook and different retail foreign exchange brokerages. The dots represent the date at which the first trader from each new brokerage joins myForexBook. Traders are not able to join the social network until their brokerage has agreed to partner with myForexBook.

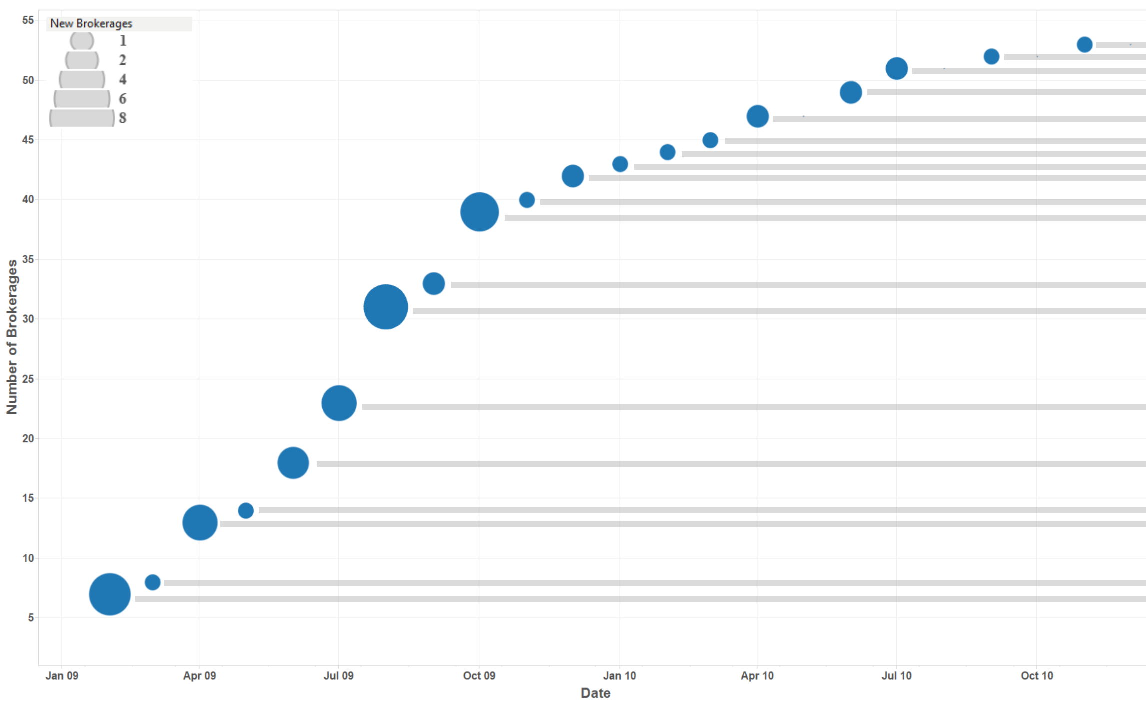
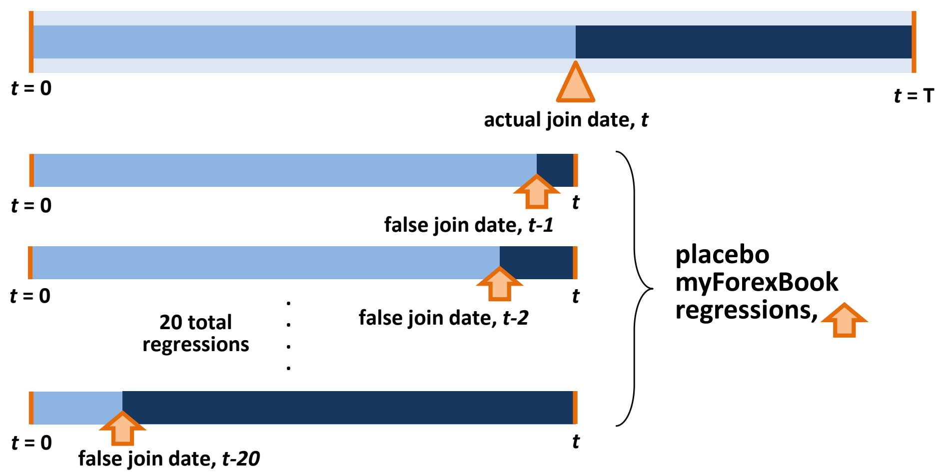


Figure 5: **Placebo Estimates of Receiver Function**

This figure (Panel II) plots kernel density estimates of the distribution of t-statistics associated with estimates of $\delta_4$ and $\delta_5$ in an exercise similar to the regression analysis described in Table 3. The treatment variable $join_{i,j,t}$ is replaced with $placebo.join_{i,j,t}$, an indicator variable that recodes the date at which a trader joins myForexBook with a false date that occurs in the prior weeks. Panel I describes how we restrict the sample and code $placebo.join_{i,j,t}$. Specifically, we redefine a trader's join date as the week before the trader actually joins, and roll back the false join date one week at a time, producing 20 total regressions.

### Panel I: An illustration of the placebo test




### Panel II: Results from placebo regressions

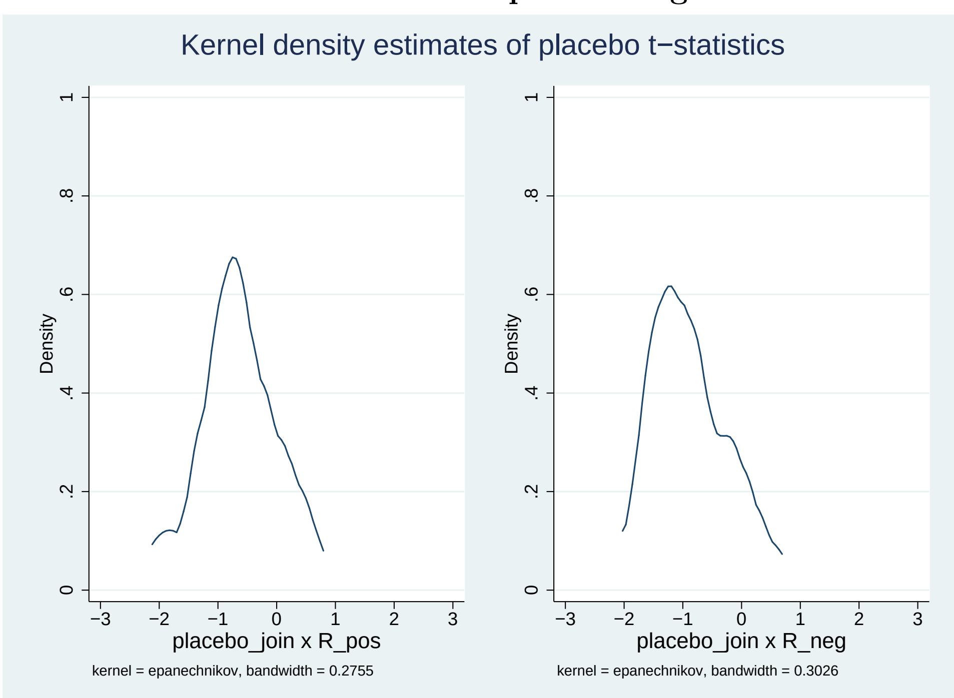


Figure 6: **Average Trading Over the Sample Period**

This figure plots a 5-week moving average of the average number of trades made in week $t$. Traders in the sample are partitioned into those that have joined myForexBook by week $t$, "post-myForexBook", and those that have not yet joined, "pre-myForexBook". Within each group of traders, each week we exclude the outer five percent of the distribution.

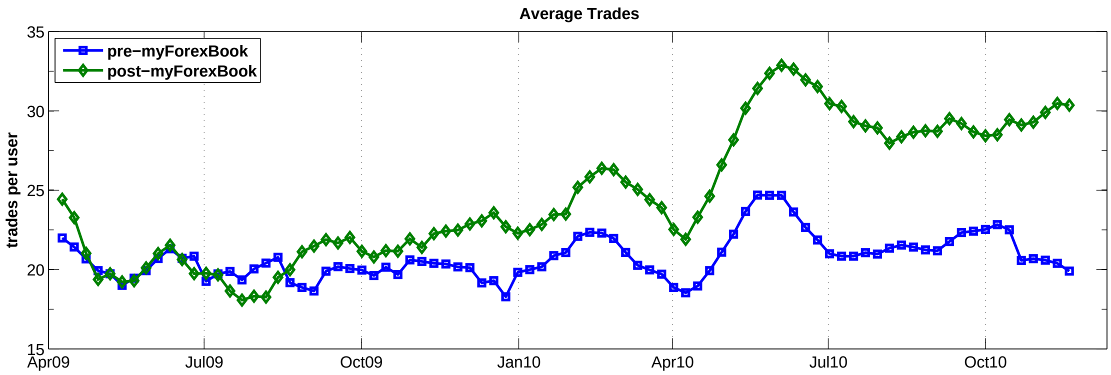


Figure 7: **Average Standard Deviation of Returns Across Traders**

This figure plots a 5-week moving average of the standard deviation of $R_{i,t}$, across traders, within week $t$. Traders in the sample are partitioned into those that have joined myForexBook by week $t$, "post-myForexBook", and those that have not yet joined, "pre-myForexBook". Within each group of traders, each week we exclude the outer five percent of the distribution.

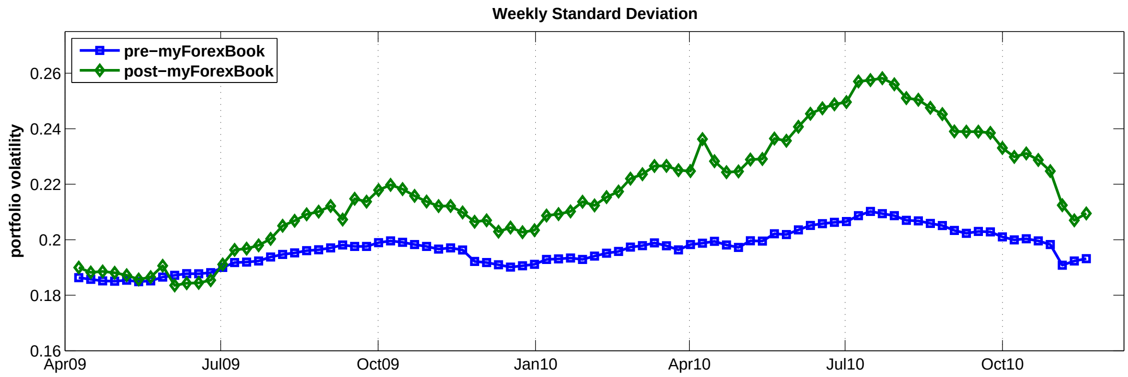


Table 1: **Summary Stats: Trading, Social Networking, and Trader Characteristics**
This table presents summary statistics from the myForexBook database. In Panel I, the weekly gross return on the trader's portfolio is $R_{i,t} = \sum_{k=1}^{s_{i,t}} p_{k,t} \cdot R_{k,t}$, where $p_{k,t}$ is the value of position $k$ when it is opened at second $t$ divided by the total opening value of all positions held by trader $i$, $R_{k,t}$ is the return on position $k$, and $s_{i,t}$ is the number of positions opened by $i$. The number of trades executed by trader $i$ in week $t$ is $trade.count_{i,t}$. In Panel II, $peer.group.size_i$, is the number of traders in $i$'s peer group $j$ as of the end of the sample period, with connections formed via friend requests initiated by one trader in a pair. The variables $sent.messages_i$ and $received.messages_i$ are the number of peer-to-peer user messages sent or received, respectively. The variables presented in Panel III are collected via a survey administered by myForexBook when traders join the network. Traders are asked their age, experience trading (years), location, and trading approach, and are able to choose from the options listed below.

| | **Panel I: Trading** | | | |
|---|---|---|---|---|
| Variable | Obs. (trader $i$, week $t$) | Mean | Median | Std. Dev. |
| $R_{i,t}$ | 111,928 | -0.028 | 0.00 | 0.20 |
| $trade.count_{i,t}$ | 111,928 | 25.54 | 11.00 | 92.20 |

| | **Panel II: Social networking** | | | |
|---|---|---|---|---|
| Variable | Obs. (trader $i$) | Mean | Median | Std. Dev. |
| $peer.group.size_i$ ($j$) | 3,117 | 21.80 | 8.00 | 33.16 |
| $sent.message.count_i$ | 3,117 | 28.17 | 17.00 | 58.00 |
| $received.message.count_i$ | 3,117 | 28.17 | 19.00 | 41.54 |

| | **Panel III: Trader characteristics** | | | | |
|---|---|---|---|---|---|
| Variable | Obs. (trader $i$) | Mean | Median | Std. Dev. | |
| $age_i$ (years) | 3,117 | 36.10 | 34.17 | 10.20 | |
| | | | | | |
| $experience_i$ (years) | 0 - 1 | 1 - 3 | 3 - 5 | 5+ | No Response |
| (fraction of traders) | 0.342 | 0.456 | 0.772 | 0.116 | 0.078 |
| | | | | | |
| $location_i$ | Asia/Pacific | Europe | U.S.A | None Specified | |
| (fraction of traders) | 0.191 | 0.443 | 0.345 | 0.021 | |
| | | | | | |
| $trading.approach_i$ | Fundamental | Momentum | News | Technical | Not Specific |
| (fraction of traders) | 0.047 | 0.057 | 0.028 | 0.634 | 0.234 |

Table 2: **A Comparison of the First and Last Traders to Join myForexBook**

This table compares the first 250, 500, and 1,000 traders to join myForexBook, $first_i$, to the last 250, 500, 1,000 traders to join, $last_i$. Panel I includes a comparison of means. Panel II estimates a Probit model in which the dependent variable $first_i$, is equal to one if a trader is among the first set of traders to join myForexBook and equal to zero if the trader is among the last to join.

**Panel I: Difference in means between first and last entrants**

| first/last network entrants | 250 | | | 500 | | | 1,000 | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | $first_i$ | $last_i$ | $t^a$ | $first_i$ | $last_i$ | $t^a$ | $first_i$ | $last_i$ | $t^a$ |
| $age_i$ | 36.384 | 35.216 | *1.31* | 35.797 | 35.406 | *0.61* | 36.488 | 36.198 | *0.63* |
| $experience_i$ | | | | | | | | | |
| 0 - 1 | 0.364 | 0.36 | *0.09* | 0.372 | 0.356 | *0.53* | 0.340 | 0.329 | *0.52* |
| 1 - 3 | 0.460 | 0.452 | *0.18* | 0.448 | 0.446 | *0.06* | 0.471 | 0.462 | *0.40* |
| 3 - 5 | 0.072 | 0.092 | *-0.81* | 0.078 | 0.086 | *-0.46* | 0.091 | 0.074 | *1.38* |
| 5 + | 0.100 | 0.080 | *0.78* | 0.096 | 0.100 | *-0.21* | 0.094 | 0.128 | *-2.42* |
| $trading.approach_i$ | | | | | | | | | |
| momentum | 0.056 | 0.048 | *0.40* | 0.066 | 0.062 | *0.26* | 0.058 | 0.058 | *0.00* |
| news | 0.036 | 0.024 | *0.79* | 0.026 | 0.030 | *-0.38* | 0.022 | 0.026 | *-0.58* |
| technical | 0.648 | 0.676 | *-0.66* | 0.622 | 0.650 | *-0.92* | 0.706 | 0.632 | *3.53* |
| not specific | 0.204 | 0.220 | *-0.44* | 0.238 | 0.210 | *1.06* | 0.175 | 0.232 | *-3.17* |
| $location_i$ | | | | | | | | | |
| Asia/Pacific | 0.192 | 0.184 | *0.23* | 0.218 | 0.218 | *0.00* | 0.176 | 0.184 | *-0.47* |
| Europe | 0.424 | 0.404 | *0.45* | 0.412 | 0.412 | *0.00* | 0.404 | 0.454 | *-2.26* |
| United States | 0.364 | 0.380 | *-0.37* | 0.348 | 0.350 | *-0.07* | 0.406 | 0.345 | *2.82* |

[a] test of equality of means among $first_i$ and $last_i$ to join myForexBook

**Panel II: Probit model estimates of being among the first entrants**

| first/last network entrants: | (a) 250 | | (b) 500 | | (c) 1,000 | |
|---|---|---|---|---|---|---|
| dep var: $first_i = 1$ | coef | (s.e.) | coef | (s.e.) | coef | (s.e.) |
| $age_i$ | 0.00691 | (0.0058) | 0.00262 | (0.0040) | 0.00139 | (0.0028) |
| $experience_i^{\dagger}$ | | | | | | |
| 0 - 1 | 0.889 | (0.69) | 0.526 | (0.45) | 0.320 | (0.39) |
| 1 - 3 | 0.908 | (0.69) | 0.511 | (0.45) | 0.284 | (0.39) |
| 3 - 5 | 0.760 | (0.72) | 0.442 | (0.47) | 0.405 | (0.40) |
| 5 + | 1.016 | (0.71) | 0.468 | (0.46) | 0.0681 | (0.40) |
| $trading.approach_i^{\ddagger}$ | | | | | | |
| momentum | -0.238 | (0.37) | 0.0184 | (0.24) | 0.162 | (0.18) |
| news | -0.0820 | (0.44) | -0.105 | (0.30) | 0.0554 | (0.23) |
| technical | -0.380 | (0.29) | -0.0406 | (0.19) | 0.256* | (0.14) |
| not specific | -0.386 | (0.30) | 0.0702 | (0.20) | -0.000292 | (0.15) |
| $location_i^{\xi}$ | | | | | | |
| Asia/Pacific | 0.215 | (0.40) | -0.123 | (0.30) | 0.0248 | (0.24) |
| Europe | 0.235 | (0.38) | -0.122 | (0.29) | -0.0193 | (0.23) |
| United States | 0.131 | (0.38) | -0.139 | (0.29) | 0.164 | (0.23) |
| *constant* | -0.976 | (0.77) | -0.458 | (0.52) | -0.571 | (0.45) |
| $N$ | 500 | | 1000 | | 2000 | |
| pseudo $R^2$ | 0.012 | | 0.0026 | | 0.012 | |

Standard errors in parentheses

* $p < 0.10$ , ** $p < 0.05$ , *** $p < 0.01$

$^{\dagger}$omitted category is *no response,* $^{\ddagger}$omitted category is *fundamental,* $^{\xi}$ omitted category is *none specified*

# Table 3: **The Receiver Function: Trading in Response to Peer Returns**

This table presents coefficients and standard errors from the following regression model estimated using OLS:

$$trade.count^{r}_{i,j,t} = \delta_1 join_{i,j,t} + \delta_2 R^{s+}_{-i,j,t} + \delta_3 R^{s-}_{-i,j,t} + \delta_4 join_{i,j,t} \times R^{s+}_{-i,j,t} + \delta_5 join_{i,j,t} \times R^{s-}_{-i,j,t} + m_t + f_i + b_i + \varepsilon_{i,j,t}.$$

The dependent variable, $trade.count^{r}_{i,j,t}$, is the number of positions opened by trader $i$ in week $t$, with the superscript $r$ indicating that $i$ is the recipient of peer effects. The indicator variable $join_{i,j,t}$ is equal to one if $i$ has joined myForexBook by week $t$. The variable $R^{s+}_{-i,j,t}$ ($R^{s-}_{-i,j,t}$) is an indicator equal to one if the average gross portfolio returns of traders in $i$'s peer group $j$ are positive (negative) in week $t$ (the superscript $s$ indicates the sender of peer effects). Peer group $j$ is defined as the set of traders that form a bilateral friendship with $i$ at any point in the sample. Trader $i$ is only able to communicate with or observe $j$'s portfolio after joining myForexBook, thus $\delta_4$ ($\delta_5$) captures the causal influence of peer effects. The variable $received.message_{i,j,t}$ is equal to one if $i$ received at least one user message in week $t$, while $average.chatter_t$ is an indicator variable equal to one if the the average number of messages per user sent within the network during week t is greater than the moving average of the previous and following five weeks. Aggregate weekly variables include, $DXY_t$, a dollar index provided by Bloomberg, and $CVIX_t$, an implied volatility index produced by Deutsche Bank. The model includes week, trader, and brokerage fixed effects, $m_t$ , $f_i$, and $b_i$ respectively. Standard errors are clustered by trader and week.

| $turnover^{r}_{i,j,t}$ | I | II | III | IV | V | VI | VII | VIII | IX | X |
|---|---|---|---|---|---|---|---|---|---|---|
| $join_{i,j,t} \times R^{s+}_{-i,j,t}$ | 4.147*** | | 5.404*** | 5.429*** | 5.697*** | 5.363** | 4.697** | 4.591** | 4.259* | 3.796* |
| | (0.88) | | (2.03) | (2.08) | (2.06) | (2.11) | (2.26) | (2.29) | (2.22) | (2.14) |
| $join_{i,j,t} \times R^{s-}_{-i,j,t}$ | 0.495 | | 0.713 | 1.785 | 0.791 | 1.414 | 0.791 | 0.774 | 0.799 | 1.674 |
| | (0.50) | | (2.12) | (2.06) | (2.21) | (2.71) | (2.21) | (2.26) | (2.64) | (3.43) |
| $join_{i,j,t}$ | | | 3.311 | 2.924 | -0.668 | 3.197 | -0.668 | -1.695 | 3.002 | 3.198 |
| | | | (2.25) | (1.96) | (2.24) | (2.21) | (2.24) | (1.20) | (2.21) | (2.21) |
| $R^{s+}_{i,j,t}$ | | -0.372 | -1.230* | -1.108 | -0.673* | -0.896* | -0.673* | -0.620 | -0.878* | -0.894* |
| | | (0.29) | (0.69) | (0.77) | (0.40) | (0.46) | (0.40) | (0.41) | (0.46) | (0.46) |
| $R^{s-}_{i,j,t}$ | | -0.128 | -1.758 | -1.781 | -1.099 | -1.470 | -1.099 | -1.086 | -1.456 | -1.468 |
| | | (0.31) | (1.44) | (1.38) | (0.87) | (1.18) | (0.87) | (0.90) | (1.18) | (1.18) |
| $join_{i,j,t} \times R^{s+}_{-i,j,t} \times received.message_{i,j,t}$ | | | | | | | | | 3.143*** | |
| | | | | | | | | | (0.91) | |
| $join_{i,j,t} \times R^{s-}_{-i,j,t} \times received.message_{i,j,t}$ | | | | | | | | | 1.763 | |
| | | | | | | | | | (1.13) | |
| $join_{i,j,t} \times R^{s+}_{-i,j,t} \times average.chatter_t$ | | | | | | | | | | 3.623*** |
| | | | | | | | | | | (0.60) |
| $join_{i,j,t} \times R^{s-}_{-i,j,t} \times average.chatter_t$ | | | | | | | | | | -0.634 |
| | | | | | | | | | | (1.88) |
| trader fixed effects | x | x | x | | x | x | x | x | x | x |
| week fixed effects | x | x | x | x | x | x | x | x | x | x |
| brokerage fixed effects | | | | x | | | | | | |
| brokerage dummies × $join_{i,j,t}$ | | | | | x | | x | x | x | x |
| brokerage dummies × week dummies | | | | | | x | | | | |
| trader dummies × $DXY_t$ | | | | | | | x | | | |
| trader dummies × $CVIX_t$ | | | | | | | | x | | |
| $N$ | 66,418 | 45,510 | 111,928 | 111,928 | 111,928 | 111,928 | 111,928 | 111,928 | 111,928 | 111,928 |
| $R^2$ | 0.0049 | 0.0047 | 0.0031 | 0.0031 | 0.031 | 0.040 | 0.031 | 0.029 | 0.040 | 0.040 |

Standard errors in parentheses

* $p < 0.10$ , ** $p < 0.05$ , *** $p < 0.01$

Table 4: **Receiver Function Alternative Specifications**

This table presents coefficients from the following regression model estimated using OLS:

$$trade.count^{r}_{i,j,t} = \delta_1 join_{i,j,t} + \delta_2 R^{s+}_{-i,j,t} + \delta_3 R^{s-}_{-i,j,t} + \delta_4 join_{i,j,t} \times R^{s+}_{-i,j,t} + \delta_5 join_{i,j,t} \times R^{s-}_{-i,j,t} + b_i \times join_{i,j,t} + m_t + f_i + \varepsilon_{i,j,t}.$$

Unless stated otherwise, peer group $j$ is defined as the set of traders that form a bilateral friendship with $i$ at any point in the sample. Standard errors are clustered by trader and week. A variety of specifications are estimated, each of which is described below.

| | | coef. estimate | | |
|---|---|---|---|---|
| Description | Specification | $join_{i,j,t} \times R^{s+}_{-i,j,t}$ | $join_{i,j,t} \times R^{s-}_{-i,j,t}$ | $join_{i,j,t}$ |
| brokerage group FE for cohorts of brokerages that agree with myForexBook in the same month | (1) | 5.47*** | 1.817 | 2.936 |
| restricts regression observations to one month before and one month after $i$ joins myForexBook | (2) | 7.031*** | 4.656 | 4.66*** |
| removes trader $i$ week $t$ observations that occur before $i$ joins myForexBook and after brokerage $b_i$ agrees with myForexBook | (3) | 4.269* | -0.368 | 7.863** |
| includes only traders $i$ who join myForexBook in the month that brokerage $b_i$ agrees with myForexBook | (4) | 2.058* | -3.77* | 2.135*** |
| uses the one week lag of peer returns, $R^{s+}_{-i,j,t-1}$ ($R^{s-}_{-i,j,t-1}$) | (5) | 3.036** | 0.055 | 3.608 |
| $R^{s+}_{-i,j,t}$ ($R^{s-}_{-i,j,t}$) are the average contemporaneous returns of any trader $j$ who messages $i$ in week $t$ † | (6) | 7.381*** | 0.447 | 2.879 |
| $R^{s+}_{-i,j,t}$ ($R^{s-}_{-i,j,t}$) are the average contemporaneous returns of any trader $j$ who messages $i$ at least once during the sample † | (7) | 7.450*** | 0.496 | 3.002 |
| $R^{s+}_{-i,j,t}$ ($R^{s-}_{-i,j,t}$) are the returns of the trader(s) in $j$ with the most total messages to trader $i$ ‡ | (8) | 6.871** | -2.622** | 2.188 |
| $R^{s+}_{-i,j,t}$ ($R^{s-}_{-i,j,t}$) are the average returns of traders in $j$ with whom $i$ forms a bilateral friendship with during the first month after joining myForexBook | (9) | 7.113** | -1.620* | 2.550 |

* $p < 0.10$ , ** $p < 0.05$ , *** $p < 0.01$

† *when $join_{i,j,t} = 0$, peer group $j$ includes the set of traders that sent at least one message to $i$ during the sample*

‡ *when traders tie for having sent the most messages to $j$, we use the average returns of those traders*

Table 5: **The Sender Function: A Trader's Returns and Contacting Others**

This table presents the probability that trader $i$ initiates communication via a user message with at least one trader during week $t$. Trader returns, $R_{i,t}$, are sorted into quartiles within the following data partitions. Trader experience, location, age, and trading approach are collected when traders join myForexBook.

| | **Probability of contacting others** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | trader experience | | trader location | | trader age | | trading approach | |
| | all traders | 0 - 3 | 4+ years | U.S.A | Europe | < 36 | 36+ | technical | all other |
| $R_{i,t}$ quartile | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| 1st | 0.156 | 0.156 | 0.156 | 0.161 | 0.153 | 0.155 | 0.157 | 0.160 | 0.148 |
| 2nd | 0.154 | 0.153 | 0.158 | 0.156 | 0.152 | 0.152 | 0.156 | 0.156 | 0.150 |
| 3rd | 0.179 | 0.177 | 0.186 | 0.181 | 0.175 | 0.182 | 0.175 | 0.183 | 0.171 |
| 4th | 0.225 | 0.214 | 0.266 | 0.225 | 0.226 | 0.215 | 0.237 | 0.220 | 0.235 |
| $N$ traders | 3,117 | 2,427 | 656 | 1,160 | 1,329 | 1,731 | 1,386 | 2,061 | 1,056 |

† cells contain the probability that $i$ sends at least one peer-to-peer message in week $t$

Table 6: **The Sender Function: An Empirical Model of Partially-Overlapping Peer Groups**

This table presents estimates of the following empirical model estimated using OLS:

$$sent.message_{i,j,t} = \beta_1 \cdot R_{i,j,t} + \beta_2 \cdot peer.\overline{R}_{-i,j,t} + \beta_3 \cdot log.peer.\overline{messages}_{-i,j,t} + \beta_4 \cdot \bar{X}_{-i,j,t} + \beta_5 \cdot X_{i,j,t} + m_t + f_i + g_{i,j} + \epsilon_{i,j,t},$$

where the dependent variable, $sent.message_{i,j,t}$, is equal to one if trader $i$ sent at least one user message in week $t$. The independent variable, $R_{i,j,t}$, is the gross returns of $i$'s portfolio in $t$, normalized to standard deviations about the mean. The natural logarithm of one plus the number of connections to $i$ is represented by $log.peer.group.size_{i,j,t}$. Peer-effects are captured by $peer.\overline{R}_{-i,j,t}$ and $log.peer.\overline{messages}_{-i,j,t}$, the average returns of peer group $j$ and the natural logarithm of one plus the average number of user messages sent by traders in $j$. Exogenous peer effects, $\bar{X}_{-i,j,t}$, include the fraction of traders in $j$ at time $t$ that have the same experience, approach, or location, which are captured by $\%.peer.experience_{-i,j,t}$, $\%.peer.trading.approach_{-i,j,t}$, and $\%.peer.location_{-i,j,t}$, respectively, as well as the natural logarithm of the average age of traders in $j$ is $log.peer.\overline{age}_{-i,j,t}$. Trader and week fixed effects are $f_i$ and $m_t$, respectively. A matrix of dummy variables, $g_{i,j}$, uses the network topology at each week $t$ and the Girvan-Newman algorithm to sort traders into nearby groups. Standard errors are clustered by trader and week.

| | *restricted.peers*[a] | | | | | *eventual.peers*[b] | | *excluded.peers*[c] | |
|---|---|---|---|---|---|---|---|---|---|
| $sent.message_{i,j,t}$ | I | II | III | IV | V | VI | VII | VIII | IX |
| $R_{i,j,t}$ (Z) | 0.0162** | 0.0164*** | 0.0151** | 0.0149** | 0.0148** | 0.0151** | 0.0152** | 0.0150** | 0.0149** |
| | (0.0070) | (0.0068) | (0.0071) | (0.0071) | (0.0070) | (0.0071) | (0.0069) | (0.0073) | (0.0073) |
| $log.peer.group.size_{i,j,t}$ | | -0.00264 | -0.00260 | -0.00254 | -0.00275* | -0.00115 | -0.00147 | 0.0290 | 0.00997 |
| | | (0.0016) | (0.0016) | (0.0016) | (0.0016) | (0.0011) | (0.0011) | (0.12) | (0.12) |
| endogenous peer effects | | | | | | | | | |
| $peer.\overline{R}_{-i,j,t}$ (Z) | | | | -0.00108 | -0.00109 | -0.000874 | -0.000886 | -0.00100 | -0.000518 |
| | | | | (0.0013) | (0.0013) | (0.0013) | (0.0013) | (0.018) | (0.018) |
| $log.peer.\overline{messages}_{-i,j,t}$ | | | | -0.00529 | -0.00454 | 0.00233 | 0.00657 | -0.00191 | -0.00183 |
| | | | | (0.0043) | (0.0043) | (0.0039) | (0.0039) | (0.0029) | (0.0029) |
| peer characteristics | | | | | | | | | |
| $log.peer.\overline{age}_{-i,j,t}$ | | | | | -0.0207 | | -0.000485 | | 0.801 |
| | | | | | (0.013) | | (0.0099) | | (0.58) |
| $\%.peer.experience_{-i,j,t}$ | | | | | 0.00701 | | -0.00855 | | 0.00292 |
| | | | | | (0.0078) | | (0.0056) | | (0.019) |
| $\%.peer.trading.approach_{-i,j,t}$ | | | | | -0.00165 | | -0.00852** | | 0.00459 |
| | | | | | (0.0062) | | (0.0043) | | (0.0092) |
| $\%.peer.location_{-i,j,t}$ | | | | | -0.0111 | | -0.00959* | | 0.000466 |
| | | | | | (0.0075) | | (0.0054) | | (0.025) |
| trader fixed effects | x | x | x | x | x | x | x | x | x |
| week fixed effects | x | x | x | x | x | x | x | x | x |
| peer-group fixed effects | | | x | x | x | x | x | x | x |
| $N$ | 66,418 | 66,418 | 66,418 | 66,418 | 66,418 | 66,418 | 66,418 | 66,418 | 66,418 |
| $R^2$ | 0.041 | 0.041 | 0.041 | 0.041 | 0.042 | 0.041 | 0.042 | 0.041 | 0.041 |

Standard errors in parentheses

* $p < 0.10$ , ** $p < 0.05$ , *** $p < 0.01$

[a] *restricted.peers includes traders j that have formed a bilateral friendship with i by week t*

[b] *eventual.peers includes traders j that eventually become friends with i, but have not done so by t*

[c] *excluded.peers includes traders j that are members of myForexBook by t, but never become friends with i*